

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 23:13:13 ; Search time 6955 Seconds

(without alignments)  
11323.105 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706

Sequence: 1 atggcgtatctatgaagcc.....cgctgacagggccttga 2706

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenBank.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pt.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_ph.\*  
24: em\_pl.\*  
25: em\_ro.\*  
26: em\_sts.\*  
27: em\_un.\*  
28: em\_vl.\*  
29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rnd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706	100.0	2706	AX420420	AX420420 Sequence
2	2701.2	99.8	3498	AY094463	AY094463 Arabidops
3	1896.8	70.1	4593	AX420419	AX420419 Sequence
C 4	1896.8	70.1	102057	AC002335	AC002335 Arabidops
C 5	797.8	29.5	801	AX420422	AX420422 Sequence
6	271.4	10.0	349	AX420423	AX420423 Sequence
7	271.4	10.0	3715	AX420424	AX420424 Sequence
C 8	148	5.5	15345	AC108755	AC108755 Oryza sat
9	108.8	4.0	137519	AT11P17	AT11P17 Arabidops
C 10	108.8	4.0	196107	ATCHRV34	ATCHRV34 Arabidops
11	87	3.2	132092	AP004011	AP004011 Oryza sat
C 12	47.6	1.8	7218	166494	166494 Sequence 14
13	45.4	1.7	161477	AC103481	AC103481 Rattus no
14	45.2	1.7	1141	AX083744	AX083744 Sequence
15	44	1.6	152079	AC094195	AC094195 Rattus no
16	42.8	1.6	256	AX261730	AX261730 Sequence
17	42.6	1.6	180418	AC106674	AC106674 Rattus no
18	42.2	1.6	1410	MUSCARGP	MUSCARGP Mus musculu
19	42.2	1.6	1595	AX305833	AX305833 Sequence
20	42.2	1.6	1595	MUSCARG	MUSCARG Mus musculu
21	42.2	1.6	1893	AK001314	AK001314 Homo sapi
22	42.2	1.6	1961	BC002343	BC002343 Homo sapi
23	42.2	1.6	1966	BC006494	BC006494 Homo sapi
24	42.2	1.6	1966	BC006516	BC006516 Homo sapi
25	42.2	1.6	2312	AK000250	AK000250 Homo sapi
26	42.2	1.6	2542	AK000221	AK000221 Homo sapi
27	42.2	1.6	2560	AK095897	AK095897 Homo sapi
28	42.2	1.6	2637	AK091742	AK091742 Homo sapi
C 29	42.2	1.6	125020	AF429315	AF429315 Homo sapi
30	42.2	1.6	1978	RATRA301	RATRA301 Rattus norv
31	41.8	1.5	191603	AL713974	AL713974 Mouse DNA
32	41.6	1.5	579	CIUC23A	CIUC23A Chinese ham
33	41.6	1.5	2142	AX401891	AX401891 Sequence
34	41.6	1.5	2229	CRUC23	CRUC23 Chinese ham
35	41.4	1.5	2052	AX305666	AX305666 Sequence
36	41.4	1.5	2052	MMSTG41	MMSTG41 Mus musculu
37	41.4	1.5	318930	AC073495	AC073495 Mus muscu
C 38	41.2	1.5	165720	AC025524	AC025524 Homo sapi
39	41	1.5	3205	AK092332	AK092332 Homo sapi
40	41	1.5	11145	HS1340191	HS1340191 Novel hum
41	41	1.5	12227	AF356524	AF356524 Homo sapi
C 42	41	1.5	143501	AC104583	AC104583 Homo sapi
C 43	41	1.5	146836	AC019317	AC019317 Homo sapi
C 44	40.8	1.5	1141	AX083744	AX083744 Sequence
45	40.6	1.5	158174	AC123065	AC123065 Mus muscu

## ALIGNMENTS

RESULT 1  
AX420420  
LOCUS AX420420  
DEFINITION Sequence 2 from Patent WO0212518.  
ACCESSION AX420420  
VERSION AX420420.1 GI:21524576  
KEYWORDS  
SOURCE  
ORGANISM  
thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 Amasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.  
Floral induction gene

JOURNAL Patent: WO 0212518-A 2 14-FEB-2002;  
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)

## FEATURES

Source

Location/Qualifiers  
1..2706  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
1..2706  
/note="unnamed protein product"

CDS

/codon\_start=1  
/protein\_id="CAD35485.1"  
/db\_xref="GI:21524577"  
/translation="MALSMKFRADSGFSGSNINWGLSPETTESDLLEFGYDI  
DRITVSSKGFATYIRHVEEVAALKALDGLANISQIIEYARPKPSLWVGII  
GPNYSKDLDEEFSEKRGILEDRLRERKTAFLIDYEMDALQASKNRGPMGSEFLR  
VDFLRQAPKREQMAGSYDNNGNMMNRPOYPHSYEDFKGDVOPSKYLMGPETPAON  
CNDEOLIHNAMLIFGEIERVSKYSPNRFALVERSEAEKSGLOGRFLNPNRIK  
MSNDELPEPDOSTFSYSGMKRSRTDMFNNDPSVSSHSTGIDGSMRPLRTNERKY  
NGAEYNDVNGKEPNMRRPSSANGTIGILPSFPGCILPSRPAQGTFRPMKSNDSWEGIDP  
AQLVRESKTRRSGSDGFTIPMGVDERSFGRGSAAPLNGPDSIDHIMWGLAKGDT  
PVCCARCPWPKGIELEKLEPVNCSATDMLAKHVAVIGELVEFVREDREDFAS  
YTEFLRLSSKRDAGVAKLDGDTLEFLVPSDFLDVLYTRQERLYGVLLKLPAY  
PVTASYROESQSNPLHYMDQARDSPANASHLYPRENTYRGAEHLTAASKRSVSP  
LRIPNNAAPQAGVSLPELLATLASILPATSOAPAEPESHQPMGSPSTVSTAHOSNL  
LYNEAPSOAKRGGPOTYHDSNOSFOOYGYOTPAQGLPPSPRYRPAASNNPYTSM  
VHGNNYOOSYNNPQSLPLNNPHNNYSKTGSSHHPIGQPMVOQIOPRASHPN  
YGPISYIQAHFQVTTNMQNLNPSOFQAKMPPAKKALEPQONOLRLQPMISDG  
OGTDEVDKNRYOSTLOFANILLQIOOKOQOOSGTPAGGP"  
46..279  
/bound\_moiety="RNA Binding Region"  
misc\_binding  
/bound\_moiety="RNA Binding Region"  
misc\_binding  
/bound\_moiety="RNA Binding Region"  
misc\_binding  
/bound\_moiety="RNA Binding Region"  
BASE COUNT 779 a 628 c 617 g 682 t  
ORIGIN

Query Match 100.0%; Score 2706; DB 6; Length 2706;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAGCCATTCAGAGCCGATGATCCGGTTTCCAGTCAAAACATCTT 60  
DB 1 ATGGCGTTATCTATGAGCCATTCAGAGCCGATGATCCGGTTTCCAGTCAAAACATCTT 60  
QY 61 TTGGTGGTACGCTTAACGCCGGAGAGACAGAGTCAATCTGACCGAGTTGTTGGAGA 120  
DB 61 TTGGTGGTACGCTTAACGCCGGAGAGACAGAGTCAATCTGACCGAGTTGTTGGAGA 120  
QY 121 TACGGGATATGATAGATCAGAGTGTATTTCTTACAGAGGCTTTGGCGTTATATATAC 180  
DB 121 TACGGGATATGATAGATCAGAGTGTATTTCTTACAGAGGCTTTGGCGTTATATATAC 180  
QY 181 AGCATGTGAGAGACAGTGTGACGCCAAGAGCGCTTCTTACAGAGCAATTTGAATGA 240  
DB 181 AGCATGTGAGAGACAGTGTGACGCCAAGAGCGCTTCTTACAGAGCAATTTGAATGA 240  
QY 241 AGTCAATTAAGATCGAATCGACAGCGCGGCAAAACCTTGTAAAGTCTATGGGGGT 300  
DB 241 AGTCAATTAAGATCGAATCGACAGCGCGGCAAAACCTTGTAAAGTCTATGGGGGT 300  
QY 301 GGAATCGGCCCTAATGTCTCAAGAGATGACCTGGAGAGAGTTCAGCAAGTTGGGAAA 360  
DB 301 GGAATCGGCCCTAATGTCTCAAGAGATGACCTGGAGAGAGTTCAGCAAGTTGGGAAA 360  
QY 361 ATGAGAGATTTTGGTTTCTCAGAGAACGACAGAGCGTTCATGATATTAATGAGATG 420  
DB 361 ATGAGAGATTTTGGTTTCTCAGAGAACGACAGAGCGTTCATGATATTAATGAGATG 420  
QY 421 GATGATGCTTTTACAGGCTAAGACATGAATGGAAGCAATGGTGGTGGTGGTGGTGG 480  
DB 421 GATGATGCTTTTACAGGCTAAGACATGAATGGAAGCAATGGTGGTGGTGGTGGTGG 480  
QY 481 GTTGAATTTCTCGGTACAGAGCGCAAAAAAGAACAAATGGGCTGGCTTACGATAC 540

DB 481 GTTGAATTTCTCGGTACAGAGCGCAAAAAAGAACAAATGGGCTGGCTTACGATAC 540  
QY 541 AGAATGCAATATGATATCAATAACCGCAGATATCTCATCATATGAAAGCTTTAAGA 600  
DB 541 AGAATGCAATATGATATCAATAACCGCAGATATCTCATCATATGAAAGCTTTAAGA 600  
QY 601 GATGTCCAGCAAGTAAAGTGTCTGTGGTGGGTTCCCTCTCTCTCTCTCTCTCTCTCT 660  
DB 601 GATGTCCAGCAAGTAAAGTGTCTGTGGTGGGTTCCCTCTCTCTCTCTCTCTCTCTCT 660  
QY 661 GATGAGCAAAATTCGCAAAATGGGATATCTCTTGGTGAATGAGAGGAGAAAAAGT 720  
DB 661 GATGAGCAAAATTCGCAAAATGGGATATCTCTTGGTGAATGAGAGGAGAAAAAGT 720  
QY 721 TACCCATCAAGAAATTTTGCACCTTGGAGATTAGAGCGGAGAGAACTCCCAATGC 780  
DB 721 TACCCATCAAGAAATTTTGCACCTTGGAGATTAGAGCGGAGAGAACTCCCAATGC 780  
QY 781 AAGGAAGCCCTACAGGGAGGATTATTCATAATCTAGATCAAAATTAATGATCTCAAC 840  
DB 781 AAGGAAGCCCTACAGGGAGGATTATTCATAATCTAGATCAAAATTAATGATCTCAAC 840  
QY 841 GATGAGTTCCTCTCTGACAGAAAGAGATCTAGTTTACTCTGTGATGAAAGCGTCAAG 900  
DB 841 GATGAGTTCCTCTCTGACAGAAAGAGATCTAGTTTACTCTGTGATGAAAGCGTCAAG 900  
QY 901 ACAATATGCTCATATATGATCTCTCATGATGATCTCTCTCTCTCTCTCTCTCTCTCT 960  
DB 901 ACAATATGCTCATATATGATCTCTCATGATGATCTCTCTCTCTCTCTCTCTCTCTCT 960  
QY 961 GGGTCTATGAGGCCCTCAGAGGTACGAATGAGCGTTCATATATGATGAGAAATCAAT 1020  
DB 961 GGGTCTATGAGGCCCTCAGAGGTACGAATGAGCGTTCATATATGATGAGAAATCAAT 1020  
QY 1021 GACGTTGTGGTAAAGAGCCAACTGGAGAGGCGCATCTGCAATATGAAATGAAATCTC 1080  
DB 1021 GACGTTGTGGTAAAGAGCCAACTGGAGAGGCGCATCTGCAATATGAAATGAAATCTC 1080  
QY 1081 CCATCTCCACAGAGCCTGGAATCTCTCCCTCCGACACAAAGTACAGAGGCCCTTATG 1140  
DB 1081 CCATCTCCACAGAGCCTGGAATCTCTCCCTCCGACACAAAGTACAGAGGCCCTTATG 1140  
QY 1141 AGGTCAAAACCCGATCTTGGGAAGATATGATCTGCTCAGTTGGTCAAGAAATGAA 1200  
DB 1141 AGGTCAAAACCCGATCTTGGGAAGATATGATCTGCTCAGTTGGTCAAGAAATGAA 1200  
QY 1201 CGAACCAGAAAGATGATGATGAGGAGGTTTACTCCAAATGGGTTGATGAGAGTCA 1260  
DB 1201 CGAACCAGAAAGATGATGATGAGGAGGTTTACTCCAAATGGGTTGATGAGAGTCA 1260  
QY 1261 TTTGGTGAAGTTCAATGCTGTAGACCTATCCGAGGCCCTGATCTGATCAATA 1320  
DB 1261 TTTGGTGAAGTTCAATGCTGTAGACCTATCCGAGGCCCTGATCTGATCAATA 1320  
QY 1321 TGGAGAGAAATGATTTCCCAAGAGTGAACCTCCGCTGTGGTCTGCTGTGTGACTAAG 1380  
DB 1321 TGGAGAGAAATGATTTCCCAAGAGTGAACCTCCGCTGTGGTCTGCTGTGTGACTAAG 1380  
QY 1381 GGAAGAGGATTAATACTAACTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1381 GGAAGAGGATTAATACTAACTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1441 AATATGCTGCTAAACATTAACGCGCTTGCATTTGATGATGATGATGATGATGATGAT 1500  
DB 1441 AATATGCTGCTAAACATTAACGCGCTTGCATTTGATGATGATGATGATGATGATGAT 1500  
QY 1501 GACAGGAGAAAGATTTTGGCTTACAGTGAATTTCTCGGATACCTTACGCAAAAGAT 1560  
DB 1501 GACAGGAGAAAGATTTTGGCTTACAGTGAATTTCTCGGATACCTTACGCAAAAGAT 1560  
QY 1561 CGGAGGAGGTTGCCAAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1620

Db	1561	CGGGGGGGTGTGGCAATATGATGATGATGATCACTTATATCTTGTCGGCCCTCATGAGAT	1620
OY	1621	TTCTTAACTGATGTACTTCCAAAGTACCCGCTCAGAAGCGGCTATATAGTGTGTTCTCAAG	1680
Db	1621	TTCTTAACTGATGTACTTCCAAAGTACCCGCTCAGAAGCGGCTATATAGTGTGTTCTCAAG	1680
OY	1661	TTACCCCCCGACGGCTTCCGTGTACAGCATCATACAGACAAGAAATCTCACTCCAATCCT	1740
Db	1661	TTACCCCCCGACGGCTTCCGTGTACAGCATCATACAGACAAGAAATCTCACTCCAATCCT	1740
OY	1741	CTGCATTTATATGATGACAAGCGCGGGAAATTCACCTGCGCATGCTATGTCACAGTTTATACCT	1800
Db	1741	CTGCATTTATATGATGACAAGCGCGGGAAATTCACCTGCGCATGCTATGTCACAGTTTATACCT	1800
OY	1801	CCTAGGGAATAATTACATTAGGGGTGACACAGACAACTTATGACAGCTGCTTCAAAACATCT	1860
Db	1801	CCTAGGGAATAATTACATTAGGGGTGACACAGACAACTTATGACAGCTGCTTCAAAACATCT	1860
OY	1861	GTTAGCGAGCTCTACAGATATCCTTAATATAGCAGCGCTCAAGCTGGGGTTAGTTTAAT	1920
Db	1861	GTTAGCGAGCTCTACAGATATCCTTAATATAGCAGCGCTCAAGCTGGGGTTAGTTTAAT	1920
OY	1921	CCGGAGCTTTTATAGCACCTCTGGCAATATTTCTCCCTGCAACTTCTCAACCGCTGGCCCT	1980
Db	1921	CCGGAGCTTTTATAGCACCTCTGGCAATATTTCTCCCTGCAACTTCTCAACCGCTGGCCCT	1980
OY	1961	GAGAGTCACCAACCTATGTACAGACCTTCAACAGATTGTTCACAGCACATCACTCCAT	2040
Db	1961	GAGAGTCACCAACCTATGTACAGACCTTCAACAGATTGTTCACAGCACATCACTCCAT	2040
OY	2041	GGACTGTACATGAGAGAAGCACCGCTTCAGCTTGGAAAGAGTCCACAACAGTTTAT	2100
Db	2041	GGACTGTACATGAGAGAAGCACCGCTTCAGCTTGGAAAGAGTCCACAACAGTTTAT	2100
OY	2101	GATCGTCGAATTCAGTCACTTCCAAACATATGGAATAGTACCTCCAGCTGGGCACTA	2160
Db	2101	GATCGTCGAATTCAGTCACTTCCAAACATATGGAATAGTACCTCCAGCTGGGCACTA	2160
OY	2161	CCTCCTCCTCCTTGCGGTTTACCCCTCAGGCTTCAACCAACCCCACTACACTAGTGGATG	2220
Db	2161	CCTCCTCCTCCTTGCGGTTTACCCCTCAGGCTTCAACCAACCCCACTACACTAGTGGATG	2220
OY	2221	GTCATATGGCAATGCAATATCCAGAGCCATCTGTAAACATGCTCAAGCTGTCCGTTA	2280
Db	2221	GTCATATGGCAATGCAATATCCAGAGCCATCTGTAAACATGCTCAAGCTGTCCGTTA	2280
OY	2281	CCAAATATGCTCAATTAATTAATTCATGATGACACTAGGGTGTGTCAAAATCATCTGTT	2340
Db	2281	CCAAATATGCTCAATTAATTAATTCATGATGACACTAGGGTGTGTCAAAATCATCTGTT	2340
OY	2341	TCTCAGCCCATGATGTCAGCAATATCAACAGACGTCATGCCCCAAACAAATATATGAT	2400
Db	2341	TCTCAGCCCATGATGTCAGCAATATCAACAGACGTCATGCCCCAAACAAATATATGAT	2400
OY	2401	CCAATTTCCAGTATTCAGCAAGCTAATTTTCATGGCTTAACAAACAAATTCAGGCAACAAC	2460
Db	2401	CCAATTTCCAGTATTCAGCAAGCTAATTTTCATGGCTTAACAAACAAATTCAGGCAACAAC	2460
OY	2461	TTTAAACCTTCCCAATTTTCAAGCGGCAATGCAACACACAGACAGATTAAGGCAAAATTTAGAG	2520
Db	2461	TTTAAACCTTCCCAATTTTCAAGCGGCAATGCAACACACAGACAGATTAAGGCAAAATTTAGAG	2520
OY	2521	CCACAAAACCAAGCACTACGATTCAGAGCTTATGATCTCTGGGGATGGTCAAGGATACACA	2580
Db	2521	CCACAAAACCAAGCACTACGATTCAGAGCTTATGATCTCTGGGGATGGTCAAGGATACACA	2580
OY	2561	GATGGGGAGTTCGATTAAGAAATCAGAAGATACAGTCAACACTACATCAATTTGACGAACCTT	2640
Db	2561	GATGGGGAGTTCGATTAAGAAATCAGAAGATACAGTCAACACTACATCAATTTGACGAACCTT	2640
OY	2641	CTTCTCCAGATACAGCAAAACACACAGCAACAGTCTTCAAGTACTCCGGCTGGACAGGG	2700
Db	2641	CTTCTCCAGATACAGCAAAACACACAGCAACAGTCTTCAAGTACTCCGGCTGGACAGGG	2700

	OY	2701 CCTGA 2706       2701 CCTGA 2706
RESULT 2	Ay094463	3498 bp mRNA linear PLN 05-MAY-2002
LOCUS	Ay094463	
DEFINITION	Arabidopsis thaliana At2g43410/Tl024.15 mRNA, complete cds.	
ACCESSION	Ay094463	
VERSION	Ay094463.1 GI:20453184	
KEYWORDS	FIL CDNA.	
SOURCE	Arabidopsis thaliana.	
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 3498)	
REFERENCE	Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamanura,Y., Yu.G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Eckert,J.R.	
AUTHORS	Arabidopsis cDNA clones unpublished 2 (bases 1 to 3498) Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamanura,Y., Yu.G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Eckert,J.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (08-ARR-2002) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
COMMENT	Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	
CDS	The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamanura,Y., Yu.G., Yu,S., Davis,R.W., Theologis,A., and Eckert,J.R.	
FEATURES	Source	
	Shinn,P. (SSF/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Eckert,J.R. (SSF/Salk) contributed equally to this work as Pis.	
	Location/Qualifiers	
	1..3498	
	/organism="Arabidopsis thaliana"	
	/db_xref="taxon:3702"	
	/chromosome="2"	
	/clone="RAFLO9-47-p10(R21966)"	
	/note="ecotype: Columbia"	
	255..2960	
	/note="unknown protein"	
	/codon_start=1	
	/product="At2g43410.Tl024.15"	
	/protein_id="AAAI9833.1"	



OY	1801	CTTGGGGAAATTTCAATTAGGGGGGACACGAAATTTGACAGTGGCTTCAAAACATCT	1880
Db	2055	CTTGGGGAAATTTCAATTAGGGGGGACACGAAATTTGACAGTGGCTTCAAAACATCT	2114
OY	1861	GTTTAGCGAGCCTCTCAGAAATACCTAAATAATGCAGCGCCTCAAGCTGGGGGTAGTTTAACT	1920
Db	2115	GTTTAGCGAGCCTCTCAGAAATACCTAAATAATGCAGCGCCTCAAGCTGGGGGTAGTTTAACT	2174
OY	1921	CCGGAGCTTTTAGCCACTCTGGCAATCTATTTCTCCGTGCAACTTCTCAACCTGTCGCCCT	1980
Db	2175	CCGGAGCTTTTAGCCACTCTGGCAATCTATTTCTCCGTGCAACTTCTCAACCTGTCGCCCT	2234
OY	1981	GAGAGTACCGCAACCTATGTCTGAGACCTTCAACAATTTGTTTCCAGAGCATCATCTCAAT	2040
Db	2235	GAGAGTACCGCAACCTATGTCTGAGACCTTCAACAATTTGTTTCCAGAGCATCATCTCAAT	2294
OY	2041	GGACTGTACAAATGGAGAGAAGCACCCCTCTCAAGCTTGAAGAAAGAGGTCCCAAAACAGTTTCAT	2100
Db	2295	GGACTGTACAAATGGAGAGAAGCACCCGTCTCAAGCTTGAAGAAAGAGGTCCCAAAACAGTTTCAT	2354
OY	2101	GATGCGCTCAAAATCAGTCAATTCCAACAATAGGAAATCAGTACACTCCAGCTGGGCAACTA	2160
Db	2355	GATGCGCTCAAAATCAGTCAATTCCAACAATAGGAAATCAGTACACTCCAGCTGGGCAACTA	2414
OY	2161	CCCTCCCTCCCTCTGGGGGTAAACCCCTCCAGCTTCAAAACCCCACTACAGTGGAGATG	2220
Db	2415	CCCTCCCTCCCTCTGGGGGTAAACCCCTCCAGCTTCAAAACCCCACTACAGTGGAGATG	2474
OY	2221	GTCCTATGGCAACATGACATAACCAAGGCCAATCTGTAAACATGCTCAGCTGTCGCGTTA	2280
Db	2475	GTCCTATGGCAACATGACATAACCAAGGCCAATCTGTAAACATGCTCAGCTGTCGCGTTA	2534
OY	2281	CCAAATATGCTCTCAATTAATTAATTTCCATGTACACTCAGGGTTCGTCAATCATCTGTT	2340
Db	2535	CCAAATATGCTCTCAATTAATTAATTTCCATGTACACTCAGGGTTCGTCAATCATCTGTT	2594
OY	2341	TTCTAGGCCAATGGTGCACGCAATTACCACACGAAAGGCGTCCATGCCAAACCCAAATATAGT	2400
Db	2595	TTCTAGGCCAATGGTGCACGCAATTACCACACGAAAGGCGTCCATGCCAAACCCAAATATAGT	2654
OY	2401	CCAATTTCCAGTTATTCAGACAAGCTAAATTTTCATGGCTTAACAAACAAATTCAGGACAGAAC	2460
Db	2655	CCAATTTCCAGTTATTCAGACAAGCTAAATTTTCATGGCTTAACAAACAAATTCAGGACAGAAC	2714
OY	2461	TTAAACCCCTTCCCAATTTCAAGCTGCATGCAACACACAGCAGATTAAGCAAAATTTAGAG	2520
Db	2715	TTAAACCCCTTCCCAATTTCAAGCTGCATGCAACACACAGCAGATTAAGCAAAATTTAGAG	2774
OY	2521	CCAAACAAACCAAGACATACCATTTCCAGACCTATATCTCTGGGGATGGTCAAGGGTACAAACA	2580
Db	2775	CCAAACAAACCAAGACATACCATTTCCAGACCTATATCTCTGGGGATGGTCAAGGGTACAAACA	2834
OY	2581	GATGGGAGAGTGCATTAAGAAATCAGAGATACAGTCAACACACTACATTAATTGCAGCAACCTT	2640
Db	2835	GATGGGAGAGTGCATTAAGAAATCAGAGATACAGTCAACACACTACATTAATTGCAGCAACCTT	2894
OY	2641	CTTCTCCAGATACAGCAAGAAACAGACAGCAACAGTCTTCAGGTACTCCGGCTGCAGAGGG	2700
Db	2895	CTTCTCCAGATACAGCAAGAAACAGACAGCAACAGTCTTCAGGTACTCCGGCTGCAGAGGG	2954
OY	2701	CCTTGA 2706	
Db	2955	CCTTGA 2960	

RESULT 3							
AX420419							
LOCUS:	AX420419		4593 bp	DNA	linear	PAT 18--JUN-2002	
DEFINITION	Sequence 1 from Patent WO0212518.						
ACCESSION	AX420419						
VERSION	AX420419.1		GI:21524575				
KEYWORDS							

SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1
AUTHORS	Amasino, R. M., Schomburg, F. M., Michaels, S. D. and Patton, D.
TITLE	Floral induction gene
JOURNAL	Patent: WO 0212518-A 1 14-FEB-2002;
FEATURES	WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
SOURCE	Location/Qualifiers 1. 4593
BASE COUNT	1275 a 958 c 978 g 1382 t
ORIGIN	/organism="Arabidopsis thaliana" /db_xref="taxon:3702"
Query Match	70.1%; Score 1896.8; DB 6; Length 4593;
Best Local Similarity	90.2%; Pred. No. 0;
Matches 2137; Conservative	0; Mismatches 2; Indels 229; Gaps 1

Query Match	70.1%	Score 1896.8;	DB 6;	Length 4593;
Best Local Similarity	90.2%;	Pred. No. 0;		
Matches 2137;	Conservative 0;	Mismatches 2;	Indels 229;	Gaps 1;

QY	568	CAGTATCCTCACTCATATGAAGACCTTTAAAGAGATGTCCAGCCAAAGTACAGTTCTGTGG	627
DB	2226	CAGTATCCTCACTCATATGAAGACCTTTAAAGAGATGTCCAGCCAAAGTACAGTTCTGTGG	2285
QY	628	ATTGGGTTCCCTCCTACTACTGTACACAAATGCATATGAGCAAAATTTCTCACAATCCGATG	687
DB	2286	ATTGGGTTCCCTCCTACTACTGTACACAAATGCATATGAGCAAAATTTCTCACAATCCGATG	2345
QY	688	ATTACCTTTGGTGAGATGCAGAGGGGTAAAAAAGTTCCCATCAAGGAATTTTGCACTTATG	747
DB	2346	ATTACCTTTGGTGAGATGCAGAGGGGTAAAAAAGTTCCCATCAAGGAATTTTGCACTTATG	2400
QY	748	GAGTTTAAAGACCGCGAGGAGCTCCCAATGCAGAGAGGCTTACAGGGAGTTATTC	807
DB	2406	GAGTTTAAAGACCGCGAGGAGCTCCCAATGCAGAGAGGCTTACAGGGAGTTATTC	2465
QY	808	AATAATCCTAACAATCAAAATTAATGTACCAAAAGATGATGGTCCCTCGACCAAGACAT	867
DB	2466	AATAATCCTAACAATCAAAATTAATGTACCAAAAGATGATGGTCCCTCGACCAAGACAT	2522
QY	868	ACTAGTTTTTACTCTGTGATGAAACGGGTCAAGAGACAGATATGTTCAATATGATCCTTCA	927
DB	2526	ACTAGTTTTTACTCTGTGATGAAACGGGTCAAGAGACAGATATGTTCAATATGATCCTTCA	2585
QY	928	TGTGTATCTTCTCCTCATTTCTACTGGAATTCCTGGGTCTATGAGGCCCTCAGAGATACG	987
DB	2586	TTTGTATCTTCTCCTCATTTCTACTGGAATTCCTGGGTCTATGAGGCCCTCAGAGATACG	2645
QY	988	AATGAGCCTTATTAATGATGAGAAATACAAATGACGTGTTGGTAAGAGCCAAACAGG	1047
DB	2646	AATGAGCCTTATTAATGATGAGAAATACAAATGACGTGTTGGTAAGAGCCAAACAGG	2705
QY	1048	AGGAGGCCATTCGCAAAATGAACATGAACTACTCCCATCTCCAAACAGACCTGGAATCCTC	1107
DB	2706	AGGAGGCCATTCGCAAAATGAACATGAACTACTCCCATCTCCAAACAGACCTGGAATCCTC	2765
QY	1108	CCATCTCCTCGACAAAGATACGAGAGCGCCCTATGAGTCAAAACCCCGATTTCTTGGGAAGA	1167
DB	2766	CCATCTCCTCGACAAAGATACGAGAGCGCCCTATGAGTCAAAACCCCGATTTCTTGGGAAGA	2825
QY	1168	TATGATCCGTGCTCAGTTGGTCAAGAAAGTAAACGAACACAGAAAGATGATCAGTGCAGC	1227
DB	2826	TATGATCCGTGCTCAGTTGGTCAAGAAAGTAAACGAACACAGAAAGATGATCAGTGCAGC	2885
QY	1228	GGTTTAACTCCAAATGCGTGTGCATGAGAGTCAATTTGGTCGAGGTTCAGTTGCTGTAGA	1287
DB	2886	GGTTTAACTCCAAATGCGTGTGCATGAGAGTCAATTTGGTCGAGGTTCAGTTGCTGTAGA	2945
QY	1288	CCATATCCGTGGCCCCCGATTCGTGATACATATGAGAGAGAAATGATGGCCAAAGGTGGA	1347
DB	2946	CCATATCCGTGGCCCCCGATTCGTGATACATATGAGAGAGAAATGATGGCCAAAGGTGGA	3005

QY 1348 ACTCCGCTGTTGCTGCTGTTGTGTAAGTATGGGAAAGGATTAAGTAAGTACT----- 1403  
 |||||  
 Db 3006 ACTCCGCTGTTGCTGCTGTTGTGTAAGTATGGGAAAGGATTAAGTAAGTACTGTA 3005  
 |||||  
 QY 1404 ----- 1403  
 Db 3066 GTACTAATTTCTAGCACTTTAACCCTTCTAGTGTCTTTCTTTTCAGAGCATTTATATAT 3125  
 |||||  
 QY 1404 ----- 1403  
 Db 3126 TTTCATTTGATTTCTGATGAGATGCTTGTATTTTCATGTTGATTTTGGCATG 3185  
 |||||  
 QY 1404 ----- 1403  
 Db 3186 TTACTGTTAGTTTCTGAGATGCTTGTATTTTCATGTTGATTTTGGCATG 3245  
 |||||  
 QY 1404 ----- 1403  
 Db 3246 CCTCAATTAAGTACTGTTGTTTCTTTTATTAATGATTTATAGGCTGAGTGTCA 3305  
 |||||  
 QY 1419 TTGTTAGCAAGAACTGATTTGAATATGCTGCTAAACATTTAGCGCTGTCATTTGATG 1478  
 |||||  
 Db 3306 TTGTTAGCAAGAACTGATTTGAATATGCTGCTAAACATTTAGCGCTGTCATTTGATG 3365  
 |||||  
 QY 1479 TGAGATCGTTTCTGTTGCTGACCAAGAGGAAAGATTTGCTGTTTACACTGAAATTC 1538  
 |||||  
 Db 3366 TGAGATCGTTTCTGTTGCTGACCAAGAGGAAAGATTTGCTGTTTACACTGAAATTC 3425  
 |||||  
 QY 1539 CCGGTACCTTAGGCAAAAGATGGGGGGGTTGCCAAATTTAGATGATGTTGATGTTG 1598  
 |||||  
 Db 3426 CCGGTACCTTAGGCAAAAGATGGGGGGGTTGCCAAATTTAGATGATGTTGATGTTG 3485  
 |||||  
 QY 1599 ATTCTGTTGCTGCTGATGATGATTTCTTAACTGATGATGATGATGATGATGATGATG 1658  
 |||||  
 Db 3486 ATTCTGTTGCTGCTGATGATGATTTCTTAACTGATGATGATGATGATGATGATGATG 3545  
 |||||  
 QY 1659 GGTATATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1718  
 |||||  
 Db 3546 GGTATATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3605  
 |||||  
 QY 1719 ACAAGATTCAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1778  
 |||||  
 Db 3606 ACAAGATTCAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3665  
 |||||  
 QY 1779 TGTAGTACAGTTTATATCTCTCTGAGGAAATTTACATTAGGGGTGACACAGATTT 1838  
 |||||  
 Db 3666 TGTAGTACAGTTTATATCTCTCTGAGGAAATTTACATTAGGGGTGACACAGATTT 3725  
 |||||  
 QY 1839 GACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1898  
 |||||  
 Db 3726 GACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3785  
 |||||  
 QY 1899 TCAAGTGTGGGTTAGTTTAACTCCGAGGCTTTTACGACCTGCTGCTGCTGCTGCTGCTG 1958  
 |||||  
 Db 3786 TCAAGTGTGGGTTAGTTTAACTCCGAGGCTTTTACGACCTGCTGCTGCTGCTGCTGCTG 3845  
 |||||  
 QY 1959 AACTTGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2018  
 |||||  
 Db 3846 AACTTGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3905  
 |||||  
 QY 2019 TTTCACAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2078  
 |||||  
 Db 3906 TTTCACAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3965  
 |||||  
 QY 2079 AAGAGTGTCAACAAAGTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2138  
 |||||  
 Db 3966 AAGAGTGTCAACAAAGTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4025  
 |||||  
 QY 2139 GTTACACTTCAGCTGGGCAACTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2198  
 |||||  
 Db 4026 GTTACACTTCAGCTGGGCAACTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4085  
 |||||  
 QY 2199 CCCCACTACACTAGTGAATGTTTCATGCAACATGCAATACAGAGCATTCCTGTTAA 2258  
 |||||

Db 4086 CCCCACTACACTAGTGAATGTTTCATGCAACATGCAATGCAAGCAATCTGTTAA 4145  
 |||||  
 QY 2259 CATGCTCAGCTGCTGCTGCTTACCAATATGCTCATATATATTTTCATGTAAGTCA 2318  
 |||||  
 Db 4146 CATGCTCAGCTGCTGCTGCTTACCAATATGCTCATATATATTTTCATGTAAGTCA 4205  
 |||||  
 QY 2319 GGGTGTGCAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2378  
 |||||  
 Db 4206 GGGTGTGCAATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4265  
 |||||  
 QY 2379 CATGCTCAGCTGCTGCTGCTTACCAATATGCTCATATATTTTCATGTAAGTCA 2438  
 |||||  
 Db 4266 CATGCTCAGCTGCTGCTGCTTACCAATATGCTCATATATTTTCATGTAAGTCA 4325  
 |||||  
 QY 2439 AACACAAATCAGGACAGACTTAAACCTTCCCAATTTCAAGTCCATGCAACACC 2498  
 |||||  
 Db 4326 AACACAAATCAGGACAGACTTAAACCTTCCCAATTTCAAGTCCATGCAACACC 4385  
 |||||  
 QY 2499 AGCAGATTAAGCAATTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2558  
 |||||  
 Db 4386 AGCAGATTAAGCAATTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4445  
 |||||  
 QY 2559 TGGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2618  
 |||||  
 Db 4446 TGGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4505  
 |||||  
 QY 2619 ACTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2678  
 |||||  
 Db 4506 ACTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4565  
 |||||  
 QY 2679 AGTACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2706  
 |||||  
 Db 4566 AGTACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4593  
 |||||

RESULT 4  
 AC002335/c 102057 bp DNA linear PLIN 11-MAR-2002  
 LOCUS Arabidopsis thaliana chromosome 2 clone T1024 map C1C10F02,  
 DEFINITION complete sequence.  
 AC002335  
 AC002335 GI:20196857  
 VERSION HTG.  
 KEYWORDS Arabidopsis thaliana.  
 ORGANISM Arabidopsis thaliana.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE  
 AUTHORS Rounsley, S.D., Teschuy, M.M., Lin, X., Ketchum, K.A., Crosby, M.L.,  
 Brandon, R.C., Spriggs, T.A., Mason, T.M., Kerlavage, A.R., Adams, M.D.,  
 Somerville, C.R. and Venter, J.C.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 102057)  
 AUTHORS Lin, X.  
 JOURNAL Direct Submission  
 REFERENCE 3 (bases 1 to 102057)  
 AUTHORS Medical Center Dr., Rockville, MD 20850, USA  
 TITLE Town, C.D. and Kaul, S.  
 JOURNAL Direct Submission  
 COMMENT Submitted (27-FEB-2002) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org  
 On Apr 18, 2002 this sequence version replaced gi:6598343.  
 FEATURES  
 source  
 1..102057  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="2"  
 /map="C1C10F02"  
 /clone="T1024"



```

misc_feature      1..16126
                  /note="overlap with BAC clone F1AB2
                  (AC004430:74158..90283)."
```

---

```

repeat_region
gene              /note="T1024.1"
                  /gene="At2g43270"
                  /complement(join(<3113..3691,3734..3924,4004..>4397))
                  /complement(join(3113..3691,3734..3924,4004..4397))
                  /gene="At2g43270"
                  /codon_start=1
                  /product="putative myb-related transcription factor"
                  /protein_id="AAB64309.2"
                  /db_xref="GI:20196861"
                  /translation="MIVVVDLLEIEIFLGLPLKSLRFTVSKQMSILSKSFAERR
                  LNVKKEKILAVGDRTLEFEGEEIOMVYLHCDIDATRSLCEGLVCIAPGMIVY
                  LNPSTQLRRFPSPNHHVSDRIREFQRELYTSPGMNMGFGDRKNGSYKVR
                  MCRSPVEKCVLVERGESELNPPRIDVGRKSRPGIILDLHTLEPINNVSPSC
                  CTMSAQNVDLDRLLATKTKNGAPEMWEIILSDVDKETSITYSTLASITINHPPE
                  SSWFVLYVSKLNLVPCDNHKLFRYNTETNELCCSSDISVISPLEDLVPLRDS
                  GHDRDYKIRFTSSCGLESKHPEAGLEFKEITTFRFVATLVGVQYRSLSKFK"
                  5148..5767
                  /gene="At2g43280"
                  /note="T1024.2; predicted by gensecan"
                  /join(<5148..5447,5552..>5767)
                  /gene="At2g43280"
                  /join(5148..5447,5552..5767)
                  /gene="At2g43280"
                  /codon_start=1
                  /product="hypothetical protein"
                  /protein_id="AAB64334.1"
                  /db_xref="GI:2289005"
                  /translation="MFYDDYSRRLGFWRMVMSCRSEKDRILLARRECKNECHQVSI
                  RKGFSYRKRPSTRECKAMIHVKYDRSGKWTYTKRKHNPGLVYSPPEARLTDE
                  KDRIOELTELELNKKRLCAAYKEQDAFAKIVYEHSNQIAKVENYVNNLKEEHELE
                  HALRSKDAI"
                  complement(5982..6953)
                  /gene="At2g43290"
                  /note="T1024.3; supported by full length cDNA:
                  Ceres:31535"
                  /complement(<5982..>6953)
                  /gene="At2g43290"
                  /complement(6214..6861)
                  /gene="At2g43290"
                  /codon_start=1
                  /product="putative calcium binding protein"
                  /protein_id="AAB64310.2"
                  /db_xref="GI:20196862"
                  /translation="MVRILYINILNSFLSLVPKKLTLPPLSPWPKTLHKNSPSP
                  STMLPSPSSSAPTRKIDPSLKRVEFMFKNGRTKTELNDLENLGIYIPDKDL
                  TOMIHKIDANGDCVDIDEFESLYSSIVDEHNDGEFEEDMKDAFVFPDGDGFFIT
                  VEELKSYMASLKGKOGKTLGCKKIMQVADGDGRVNYKEFLQMKGGSSSN"
                  complement(6998..7029)
                  /rpt_family="AT_rich"
                  /complement(7486..7547)
                  /rpt_family="AT_rich"
                  /complement(7628..7660)
                  /rpt_family="AT_rich"
                  /complement(10479..10550)
                  /gene="At2g43300"
                  /complement(10479..10550)
                  /gene="At2g43300"
                  /product="tRNA-Thr"
                  11012..11413
                  /gene="At2g43310"
                  /note="T1024.5; predicted by gensecan"
                  <11012..>11413
                  /gene="At2g43310"
                  11012..11413
                  /gene="At2g43310"
                  /codon_start=1
```

---

```

repeat_region
gene              /product="hypothetical protein"
                  /protein_id="AAB64333.1"
                  /db_xref="GI:2289004"
                  /translation="MNGTNRHHLQLVLSCRKITYAOTPOGSSITIAMSSDEEFLA
                  KSRANLRRPHSHNFMDSKTASRGEKLGRLRGLADVVSIDADEEISRPRIHRRKV
                  LPLFDSVRRRTGIRVDGTEQLNDISVPILPN"
                  complement(11649..14833)
                  /gene="At2g43320"
                  /note="T1024.6; predicted by gensecan; supported by cDNA:
                  g1_14517475_5b_A039573.1"
                  /complement(join(<11649..11950,12040..12519,12610..12654,
                  13002..13102,13191..13257,13742..13804,13936..14107,
                  14566..14833))
                  /gene="At2g43320"
                  /complement(join(11780..11950,12040..12519,12610..12654,
                  13002..13102,13191..13257,13742..13804,13936..14064))
                  /gene="At2g43320"
                  /codon_start=1
                  /product="expressed protein"
                  /protein_id="AAB64311.2"
                  /db_xref="GI:20196858"
                  /translation="MRAPSLLAQCLPELLPDRGCVALSSEKIDLEPTPAVEITPSKT
                  VAHRRYSGENLDALGLQVFKGVADIIIGLSETPAPLNKESLKSWESSVAVNLV
                  KNEIRDQLSFRGKRVLELGCNFGVPGIFAFCLKGAASSVHODLSAETIRCTIPNVLA
                  NLEQARDROSROPESPLTPSRQATASAVRFAGMEELSTVLSIRFDVLEPSIPANN
                  LSPSEDFMDGSSODGSTIGDPDPSRSRKSLSKASAMERANETGEGEYGVITL
                  TELPYSTLSLKLKLSLTKCIRPPYGMVLAAKQYGFNSGAKHLNLVDEETILGA
                  HLVEETDRDIWKEFLK"
                  16041..18948
                  /gene="At2g43330"
                  /note="T1024.7"
                  /join(<16041..16188,16408..16514,16563..16898,17700..17993,
                  18080..18948)
                  /gene="At2g43330"
                  /join(16041..16188,16408..16514,16563..16898,17700..17993,
                  18080..18760)
                  /gene="At2g43330"
                  /codon_start=1
                  /product="putative membrane transporter"
                  /protein_id="AAB64332.1"
                  /db_xref="GI:2289003"
                  /translation="MTLTIPNAPGSSGYLDMFPRRMSYFGNSYILGLTVYAGIGLL
                  GYDTGVISGALVYIKDDEFVYKOSFLOYVAVSFSKLETTIVSMALVGMAMPAAY
                  GGMTNDVYGRKATLFDVYFAAGAIYMAAPVYVILSGILGLGVAVSAPVAPV
                  IAEASPESEVKGVLSTVNLMTGQFLSTLVNSAFTQVPGYRMHLYSGVPAYQIT
                  LMLFMPESPWLFEKRRKAEALIVLARTYDISRLDELDLSAAEEBKOKKRVGL
                  DVFRSKELRLAFLAGAGIQAFQFTGNTVWYSPITYVQNAFGSNLALFLSLIVA
                  MNAAGTVVGIYFIDHCGRKLALSILFGVITSLILSVSEFKQSETSDDGLYGMIAV
                  LGLALIVFPAFGMGPVPMVYSEIYPOOVYIGCGSAYATNMTSNLIVQOTFLTIE
                  AAGCGMTFLIAGIYAVLAVIFVYFVETQGLTSEVQIWKERAYGNISGSSSS
                  NMEGLEBOGSS"
                  complement(21849..21900)
                  /rpt_family="AT_rich"
                  22675..23510
                  /gene="At2g43340"
                  /note="T1024.8"
                  /join(<22675..23164,23243..23304,23325..23510)
                  /gene="At2g43340"
                  /join(22675..23164,23243..23304,23325..23498)
                  /gene="At2g43340"
                  /codon_start=1
                  /product="unknown protein"
                  /protein_id="AAB64331.1"
                  /db_xref="GI:2289002"
                  /translation="MCDSDFLSRCSYEADVDSDEISTSSFSQSYSESEERINNG
                  FGVYOSKQFKLEKKSNVLEGVYVDSAYANDDLKRSITLDDLELGLGCVLDFG
                  FNYEIEPEICNTLPALFALCYSMOKFIDDDHNNSSPEKSSVLDSPVSLSMIT
                  SSPGDNDPVYAKRLKFAQAVACTMRCGFQPKABESMEBEDEFTQFAGVSKLML
                  TCLKTTIVLELAHTVCTKRLMT"
                  complement(23549..25505)
                  /gene="At2g43350"
                  /note="T1024.9; supported by full length cDNA:
                  Ceres:30737"
```

mRNA	complement(join(23549..23749,24086..24133,24215..24382,24465..24583,24673..24734,24830..24906,25278..25505))	
	/gene="At2g43350"	
CDS	complement(join(24101..24133,24215..24382,24465..24583,24673..24734,24830..24906,25278..25439))	
	/gene="At2g43350"	
gene	/product="putative glutathione peroxidase"	
	/protein_id="AAB64335.1"	
mRNA	/db_xref="GI:2289006"	
	/translation="MRRSRVNVORATSKKRTLELGVAFVYLYRPSPEVDSSTINISVDIEGKDVLSKFTGKVLIVNVAISKGLTHGNKENMILAKTKQGEELIAPCNQSQSPGSMKEIKETVCNIFKAPFPIEDKIEGNKNCPLNFKEQGGLFDALKNMFAELVDROGNVVDYRAPTTSPLIEKIDYKLIASA"	
gene	complement(25700..28116)	
	/gene="At2g43360"	
mRNA	/note="TI024.10; supported by full length cDNA: Ceres:42038"	
	complement(join(<25700..26123,26213..26352,26523..26728,	
Query Match	70.1%; Score 1896.8; DB 8; Length 102057;	
	Best Local Similarity 90.2%; Pred. No. 0;	
Matches 2137; Conservative 0; Mismatches 2; Indels 229; Gaps 1;		
QY	568 CAGTATCTCAGTCATATGAAGACTTTAAAGAGATGCCCAAGTAAGTTCTGTGG	627
Db	43670 CAGTATCTCAGTCATATGAAGACTTTAAAGAGATGCCCAAGTAAGTTCTGTGG	43611
QY	628 ATTGGGTTCCCTCTACTGCTACACAAATGCAATGATGAGCAAAATTCGCACATGGCATG	687
Db	43610 ATTGGGTTCCCTCTACTGCTACACAAATGCAATGATGAGCAAAATTCGCACATGGCATG	43551
QY	688 ATACTCTTTGGTGAATCGAGAGGGTAAAAAGTTACCATCAAGAAATTTTGCATTGTG	747
Db	43550 ATACTCTTTGGTGAATCGAGAGGGTAAAAAGTTACCATCAAGAAATTTTGCATTGTG	43491
QY	748 GAGTTTAGAGCGCGGAGAAAGCTCGCCATGCAAGGAAGGCTTACAGGGAGATTATTC	807
Db	43490 GAGTTTAGAGCGCGGAGAAAGCTCGCCATGCAAGGAAGGCTTACAGGGAGATTATTC	43431
QY	808 AATTAATCTGAATCAAAATTTATGTAATCAAAAGATGTTGCCCTCTGAGCAAGCAT	867
Db	43430 AATTAATCTGAATCAAAATTTATGTAATCAAAAGATGTTGCCCTCTGAGCAAGCAT	43371
QY	868 ACTAGTTTACTCTGTGATGAACGCTCAAGACAGATATGTTCAATAATGATCCTTCA	927
Db	43370 ACTAGTTTACTCTGTGATGAACGCTCAAGACAGATATGTTCAATAATGATCCTTCA	43311
QY	928 TGTGATCTCTCTCTATCTACTGGAATTCCTGGGCTATGAGGCCCTCAAGGTAAC	987
Db	43310 TGTGATCTCTCTCTATCTACTGGAATTCCTGGGCTATGAGGCCCTCAAGGTAAC	43251
QY	988 AATGAGCGTTCAATATATGATGTCAGATACATGAGTTGTTGTAAGAGCCAAACTGG	1047
Db	43250 AATGAGCGTTCAATATATGATGTCAGATACATGAGTTGTTGTAAGAGCCAAACTGG	43191
QY	1048 AAGAGCCATCTGCAATGGAATGGAATTAATCTCCATCTCAACAGAGCTGGAATCTTC	1107
Db	43190 AAGAGCCATCTGCAATGGAATGGAATTAATCTCCATCTCAACAGAGCTGGAATCTTC	43131
QY	1108 CCATCTCTGCAACAGTACGAGCGCCCTATGAGTCAAAACCCGATCTTGGGAAGGA	1167
Db	43130 CCATCTCTGCAACAGTACGAGCGCCCTATGAGTCAAAACCCGATCTTGGGAAGGA	43071
QY	1168 TATGATCTGCTCAGTTGTCAGAGAAATTAACAGAACAGAGATGATCAGTGAGAC	1227
Db	43070 TATGATCTGCTCAGTTGTCAGAGAAATTAACAGAACAGAGATGATCAGTGAGAC	43011
QY	1228 GGTTTACTCCATGGGTGTGATGAGAGTCAATTTGGTCAAGTTCAATTGCTGTAGA	1287
Db	43010 GGTTTACTCCATGGGTGTGATGAGAGTCAATTTGGTCAAGTTCAATTGCTGTAGA	42951
QY	1288 CCTATCCGTGGCCCCCTGATTTCTGATCATATGAGAGGAATGATTTCCCAAGGTGGA	1347
Db	42950 CCTATCCGTGGCCCCCTGATTTCTGATCATATGAGAGGAATGATTTCCCAAGGTGGA	42891
QY	1348 ACTCCGTCGTGTGTGCTGCTGTGTAACCATGAGGAAAGGGATTTGAACTAACT---	1403
Db	42890 ACTCCGTCGTGTGTGCTGCTGTGTAACCATGAGGAAAGGGATTTGAACTAACTGTGA	42831
QY	1404 -----	1403
Db	42830 GTACTAATTTCTAGCACTTTAAACCTCTAGTGTTCCTTTTCAGAGCATTTATATAT	42771
QY	1404 -----	1403
Db	42770 TTTCATTTCAATCTCGATGAGAGTAACATTTATTATAGATGATTTTATTATTCTA	42711
QY	1404 -----	1403
Db	42710 TTACTGTTTATGTTTCTGAGATGCTGATTTTCATGAGTGTGATTCATTTTGGCATTTG	42651
QY	1404 -----	1403
Db	42650 CCCTCAATTTACTGACTTTGTTTTTTTAAATATGATTTATAGGCTGAGGTCGTCAA	42591
QY	1419 TTGTTAGCAAGAACTGATTTGAATATGCTGCTAAACATTTACCGCTTGCCATTGGATG	1478
Db	42590 TTGTTAGCAAGAACTGATTTGAATATGCTGCTAAACATTTACCGCTTGCCATTGGATG	42531
QY	1479 TGAATCGTTTTTTTGGTACACAGAGGAAAGATTTTGGCTCTTACAGTAATTTCT	1538
Db	42530 TGAATCGTTTTTTTGGTACACAGAGGAAAGATTTTGGCTCTTACAGTAATTTCT	42471
QY	1539 CCGGTACCTTACGTCAAAAGATGGGCGGGTGTGGCAATTTAGATGATGATCAACTTT	1598
Db	42470 CCGGTACCTTACGTCAAAAGATGGGCGGGTGTGGCAATTTAGATGATGATCAACTTT	42411
QY	1599 ATTCTGGTGCCTCCATCAGATTTCTTAATCTGATGATCTCAAGTACCCGTCAGAACG	1658
Db	42410 ATTCTGGTGCCTCCATCAGATTTCTTAATCTGATGATCTCAAGTACCCGTCAGAACG	42351
QY	1659 GCTATATGGTGTGTTCTCAAGTTTACCCGCGAGCGGTTCCGTGACAGATCAATACAG	1718
Db	42350 GCTATATGGTGTGTTCTCAAGTTTACCCGCGAGCGGTTCCGTGACAGATCAATACAG	42291
QY	1719 ACAAGATCTCAGTCCAAATCTCTGCAATATATGATGATCAAGCCCGGATTCACCTGCCAA	1778
Db	42290 ACAAGATCTCAGTCCAAATCTCTGCAATATATGATGATCAAGCCCGGATTCACCTGCCAA	42231
QY	1779 TGCTAGTCAAGTTTATATCTCTAGGAAAAATTCATTAGGGGTGACACGAACATTT	1838
Db	42230 TGCTAGTCAAGTTTATATCTCTAGGAAAAATTCATTAGGGGTGACACGAACATTT	42171
QY	1839 GACAGCTGCTCAAAACATCTGTTAGCAGAGCTCCAGAAATACCTAATATATGACAGCGC	1898
Db	42170 GACAGCTGCTCAAAACATCTGTTAGCAGAGCTCCAGAAATACCTAATATATGACAGCGC	42111
QY	1899 TCAAGCTGGGGTATGTTAACTCCGAGCTTTTACGACTGTGGCATCTATTCCTCCGTC	1958
Db	42110 TCAAGCTGGGGTATGTTAACTCCGAGCTTTTACGACTGTGGCATCTATTCCTCCGTC	42051
QY	1959 AACTTCTCAACGCTGGCGCCCTGAGAGTACCAACCTATGTCAGACCTTCAACAGTGT	2018
Db	42050 AACTTCTCAACGCTGGCGCCCTGAGAGTACCAACCTATGTCAGACCTTCAACAGTGT	41991
QY	2019 TTCCACAGCATCAAGTCCAAATGAGTGTACATGAGAGACAGCGTCTCAAGCTTGGAA	2078
Db	41990 TTCCACAGCATCAAGTCCAAATGAGTGTACATGAGAGACAGCGTCTCAAGCTTGGAA	41931
QY	2079 AAGAGTCCACAAACAGTTTCATGATGCTGCTCAATATCATGATCTTCCAAATATGAGAAATCA	2138
Db	41930 AAGAGTCCACAAACAGTTTCATGATGCTGCTCAATATCATGATCTTCCAAATATGAGAAATCA	41871
QY	2139 GTACACTCCAGCTGGGCAACTACCTCTCTCGGTTTACCTCCAGCTTCAACAA	2198



Db 41870 GTACACTCAGCTGGGCAACTACCTCCTCTCGGCTTACCCTCCAGCTCAACAA 41811  
Qy 2199 CCCCACTACACTGTGGAAATGTCATGCGCAACATGCAATACAGAGCCATCTGTTAA 2258  
Db 41810 CCCCACTACACTGTGGAAATGTCATGCGCAACATGCAATACAGAGCCATCTGTTAA 41751  
Qy 2259 CATGCTCAGCTGATCGCTGTTACCAAAATATGCTCATATATATTTATTCATGTAACCA 2318  
Db 41750 CATGCTCAGCTGATCGCTGTTACCAAAATATGCTCATATATATTTATTCATGTAACCA 41651  
Qy 2319 GGGTTCTGCAATATCATCTGTTTCTCAGCCCATGCTCAGCAATACCAACAGAAAGCTC 2378  
Db 41690 GGGTTCTGCAATATCATCTGTTTCTCAGCCCATGCTCAGCAATACCAACAGAAAGCTC 41631  
Qy 2379 CATGCCAAACCAAACTATGTTCCATTCGAAGTTATAGCAAGCTAATTTTTCATGCGGT 2438  
Db 41630 CATGCCAAACCAAACTATGTTCCATTCGAAGTTATAGCAAGCTAATTTTTCATGCGGT 41571  
Qy 2439 AACAACTATGAGCAGACAGAACTTAAACCTTCCCAATTTCAAGCTGCATGCAACAGC 2498  
Db 41570 AACAACTATGAGCAGACAGAACTTAAACCTTCCCAATTTCAAGCTGCATGCAACAGC 41511  
Qy 2499 AGCAGATTAAGCAATTTTATAGCCCAAAACCAAGCACTAGATTGCAAGCTATGATCTC 2558  
Db 41510 AGCAGATTAAGCAATTTTATAGCCCAAAACCAAGCACTAGATTGCAAGCTATGATCTC 41451  
Qy 2559 TGGGATGCTCAGGCTACACAGATGGGAGGTGATTAAGATCAGATACAGATACAGTCAAC 2618  
Db 41450 TGGGATGCTCAGGCTACACAGATGGGAGGTGATTAAGATCAGATACAGATACAGTCAAC 41391  
Qy 2619 ACTCAATTTGACAGAAACCTTCTTCTCAGATACAGAGAAACAGCAACAGCACTCTC 2678  
Db 41390 ACTCAATTTGACAGAAACCTTCTTCTCAGATACAGAGAAACAGCAACAGCACTCTC 41331  
Qy 2679 AGGTACTCCGGCTGGACAGGGGCTTGA 2706  
Db 41330 AGGTACTCCGGCTGGACAGGGGCTTGA 41303

RESULT 5  
AX420422/c 801 bp DNA linear PAT 18-JUN-2002  
LOCUS AX420422  
DEFINITION Sequence 4 from Patent WO0212518.  
ACCESSION AX420422  
VERSION AX420422.1 GI:21524578  
KEYWORDS  
SOURCE  
ORGANISM  
thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 Amasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.  
TITLE Floral induction gene  
JOURNAL Patent: WO 0212518-A 4 14-FEB-2002;  
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)  
FEATURES  
Source  
Location/Qualifiers  
1..801  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/note="FPA antisense fragment"

BASE COUNT 206 a 207 c 166 g 222 t  
ORIGIN  
Query Match 29.5%; Score 797.8; DB 6; Length 801;  
Best Local Similarity 99.8%; Pred. No. 2.1e-220;  
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 587 AAGACTTTAAAGAGATGCTCAGCAAGTAAAGTTCTGTGATGGTTCCCTCTACTG 646  
Db 801 AAGACTTTAAAGAGATGCTCAGCAAGTAAAGTTCTGTGATGGTTCCCTCTACTG 742  
Qy 647 CTACACATGCAATGATGAGCAAAATTTCTGCACAAATGCATGATCTTTGGTGATGCG 706

Db 741 CTACACATGCAATGATGAGCAAAATTTCTGCACAAATGCATGATCTTTGGTGATGCG 682  
Qy 707 AGAGGGTAAAAAGTTACCATCAAGCAATTTTGCACCTGTGTGAGTTAGAGAGCGGAGG 766  
Db 681 AGAGGGTAAAAAGTTACCATCAAGCAATTTTGCACCTGTGTGAGTTAGAGAGCGGAGG 622  
Qy 767 AAGCTCGCCATGCAAGCAAGAGGCTCAGAGGGAGGTTATTCATATATCTCAATCAAA 826  
Db 621 AAGCTCGCCATGCAAGCAAGAGGCTCAGAGGGAGGTTATTCATATATCTCAATCAAA 562  
Qy 827 TTATGTTCAAAACATGATGCTCCTGAGCAAGCAATGATGATTTTACTGTGTA 886  
Db 561 TTATGTTCAAAACATGATGCTCCTGAGCAAGCAATGATGATTTTACTGTGTA 502  
Qy 887 TGAACGGTCAAGCAAGATATGTCATATATGATCTTCAATGATGATCTCTCTCAT 946  
Db 501 TGAACGGTCAAGCAAGATATGTCATATATGATCTTCAATGATGATCTCTCTCAT 442  
Qy 947 CTACTGGAATTCCTGGTCTATGAGGCCCTCAGAGTACGAATGAGGTTCAATTAATG 1006  
Db 441 CTACTGGAATTCCTGGTCTATGAGGCCCTCAGAGTACGAATGAGGTTCAATTAATG 382  
Qy 1007 GTGCAATATACATGACGTTGTTGGTAAAGAGCCAAACCTGAGAGGCCATCTGCAATG 1066  
Db 381 GTGCAATATACATGACGTTGTTGGTAAAGAGCCAAACCTGAGAGGCCATCTGCAATG 322  
Qy 1067 GAACGTGAATCTCCATCTCCACAGACAGCTGGAATCTCCATCTCTGACAGAGTA 1126  
Db 321 GAACGTGAATCTCCATCTCCACAGACAGCTGGAATCTCCATCTCTGACAGAGTA 262  
Qy 1127 CGAGCGCCCTATGAGGTCAAACCCGATTTCTGGGAAGATATATCTGCTCAGTTGG 1186  
Db 261 CGAGCGCCCTATGAGGTCAAACCCGATTTCTGGGAAGATATATCTGCTCAGTTGG 202  
Qy 1187 TCAGAGAAAGTAAAGCAACAGAGAGATGATGATGATGATGATGATGATGATGATG 1246  
Db 201 TCAGAGAAAGTAAAGCAACAGAGAGATGATGATGATGATGATGATGATGATGATG 142  
Qy 1247 TCGATGAGAGTCAATTTGGTCGAGGTCGATGTCGATGATGATGATGATGATGATG 1306  
Db 141 TCGATGAGAGTCAATTTGGTCGAGGTCGATGTCGATGATGATGATGATGATGATG 82  
Qy 1307 ATTCTGATCACAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1366  
Db 81 ATTCTGATCACAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 22  
Qy 1367 GTTGTACTATGAGGAAAG 1387  
Db 21 GTTGTACTATGAGGAAAG 1

RESULT 6  
AX420423 349 bp DNA linear PAT 18-JUN-2002  
LOCUS AX420423  
DEFINITION Sequence 5 from Patent WO0212518.  
ACCESSION AX420423  
VERSION AX420423.1 GI:21524579  
KEYWORDS  
SOURCE  
ORGANISM  
thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 Amasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.  
TITLE Floral induction gene  
JOURNAL Patent: WO 0212518-A 5 14-FEB-2002;  
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)  
FEATURES  
Source  
Location/Qualifiers  
1..349  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"

BASE COUNT 92 a 68 c 81 g 108 t  
 ORIGIN /note="Portion of FPA coding region"

Query Match 10.0%; Score 271.4; DB 6; Length 349;  
 Best Local Similarity 94.6%; Pred. No. 3.1e-67;  
 Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATCCGGTTCCAGTCAAAATCTT 60  
 Db 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATCCGGTTCCAGTCAAAATCTT 60

QY 61 TGGGTGGTACGCTTACGCGGAGACAGAGTCAATGATCGACGAGTTGTTGAAGA 120  
 Db 61 TGGGTGGTACGCTTACGCGGAGACAGAGTCAATGATCGACGAGTTGTTGAAGA 120

QY 121 TACGGCGATTTGATTAATATCAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180  
 Db 121 TACGGCGATTTGATTAATATCAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180

QY 181 AGACATGTGAGAGAGCAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240  
 Db 181 AGACATGTGAGAGAGCAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240

QY 241 AGTCAATTTAAGATCGAATACGACGACCGGTTGTTCTTATCTATCTTCGTTTG 297  
 Db 241 AGTCAATTTAAGATCGAATACGACGACCGGTTGTTCTTATCTATCTTCGTTTG 297

RESULT 7  
 AX420424 3715 bp DNA linear PAT 18-JUN-2002  
 LOCUS  
 DEFINITION Sequence 6 from Patent WO0212518.  
 ACCESSION AX420424  
 VERSION AX420424.1 GI:21524580  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1  
 Anasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.  
 JOURNAL  
 TITLE  
 PATENT: WO 0212518-A 6 14-FEB-2002;  
 WISCONSIN ALUMNI RESEARCH FOUNDATION (US)  
 FEATURES  
 source  
 1. 3715  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /note="FPA promoter plus intron"

BASE COUNT 1055 a 728 c 711 g 1221 t  
 ORIGIN

Query Match 10.0%; Score 271.4; DB 6; Length 3715;  
 Best Local Similarity 94.6%; Pred. No. 3.3e-67;  
 Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATCCGGTTCCAGTCAAAATCTT 60  
 Db 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATCCGGTTCCAGTCAAAATCTT 60

QY 1832 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATCCGGTTCCAGTCAAAATCTT 1891  
 Db 1832 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATCCGGTTCCAGTCAAAATCTT 1891

QY 61 TGGGTGGTACGCTTACGCGGAGACAGAGTCAATGATCGACGAGTTGTTGAAGA 120  
 Db 61 TGGGTGGTACGCTTACGCGGAGACAGAGTCAATGATCGACGAGTTGTTGAAGA 120

QY 1892 TGGGTGGTACGCTTACGCGGAGACAGAGTCAATGATCGACGAGTTGTTGAAGA 1951  
 Db 1892 TGGGTGGTACGCTTACGCGGAGACAGAGTCAATGATCGACGAGTTGTTGAAGA 1951

QY 121 TACGGCGATTTGATTAATATCAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180  
 Db 121 TACGGCGATTTGATTAATATCAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180

QY 1952 TACGGCGATTTGATTAATATCAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 2011  
 Db 1952 TACGGCGATTTGATTAATATCAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 2011

QY 181 AGACATGTGAGAGAGCAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240  
 Db 181 AGACATGTGAGAGAGCAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240

QY 2012 AGACATGTGAGAGAGCAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 2071  
 Db 2012 AGACATGTGAGAGAGCAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 2071

QY 241 AGTCAATTTAAGATCGAATACGACGACCGGTTGTTCTTATCTATCTTCGTTTG 297  
 Db 2072 AGTCAATTTAAGATCGAATACGACGACCGGTTGTTCTTATCTATCTTCGTTTG 2128

RESULT 8  
 AC108755/c 154345 bp DNA linear HTG 31-JAN-2002  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone OSJNBa0028C22, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 2 ordered pieces.  
 AC108755  
 AC108755.1 GI:18449960  
 HTG; HTGS\_PHASE2.  
 Oryza sativa (japonica cultivar-group).  
 Oryza sativa (japonica cultivar-group).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoideae; Oryza.  
 1 (bases 1 to 154345)  
 Yoon, U.-H., Hahn, J.-H., Lee, J.-S., Yun, D.-W., Lee, M.-C., Eun, M.Y. and Kim, H.-I.  
 Oryza sativa BAC OSJNBa0028C22 genomic sequence  
 Unpublished  
 2 (bases 1 to 154345)  
 Hahn, J.-H. and Kim, H.-I.  
 Direct Submission  
 Submitted (31-JAN-2002) Rice Genome Sequencing Project, National Institute of Agricultural Science and Technology (NIAST), RDA, 249 Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@nrd.90.kr, Tel: 82-31-290-0309, Fax: 82-31-290-0308)  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces \* is believed to be correct as given, however the sizes \* of the gaps between them are based on estimates that have \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and \* the accession number will be preserved.  
 1 130145: contig of 130145 bp in length  
 \* 130146 130245: gap of unknown length  
 \* 130246 154345: contig of 24100 bp in length.  
 Location/Qualifiers  
 1. 154345  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /variety="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="9"  
 /clone="OSJNBa0028C22"

BASE COUNT 43614 a 31420 c 31651 g 45627 t 2033 others  
 ORIGIN

Query Match 5.5%; Score 148; DB 2; Length 154345;  
 Best Local Similarity 67.5%; Pred. No. 2.8e-31;  
 Matches 208; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1405 CCTGAGGTGTCATTTGTCAGCAAGACGATTTAAATATGTCCTAAACATTTGCGCC 1464  
 Db 144249 CCGGATGTTGTTATTTGTCCTGTAACCTGAGCTGATATGCTGCAAAAGCATTTAGA 144190

QY 1465 GTTGCATTTGATGAGATCGTTTTTTTCGTAACGACAGAGGAAAGATTTTGGCTCT 1524  
 Db 144189 GACGCTTCAGGGTTGATATCGCTCTCTTCCGAGATGAGATGACTTTGTTCT 144130

QY 1525 TACACTGAATTTCTCGGTACCTTACCTAAGATCGGGCGGCTGTTCCAAATTTAGAT 1584  
 Db 144129 TACACTGAATTTTGGCTTACCTTAAGCTCAAAAGCTCGGGAGGTTGTAAGGTTGAT 144070

QY 1585 GATGTACAACTTTATTTGTTGCTCCATCAGATTTTCTTAAGTATGATCTACCAAGTG 1644  
 Db 144069 GGGGGACACTATTATTTTGGTCCACCTTCGATTTTTCGAAATGTTTTCGACAGTT 144010



OY	1464	CCTTCCCAATGGATGAGATCGTTTTCGTACACAGCAGGGAAGAATTTGC	1523
DB	134307	CCAATCATCTAAAGCATGGGTGTTTTCTTTGTTCTTGGAAGATGCTAATCTGTT	134366
OY	1524	TTACACGAAATTTCTCCGTACTTAAGCTCAAAGATCGGGGGGTGCCAAATTAGA	1583
DB	134367	TTATGATGATTTATGATCATATCTGGAGSAGAAAGCAGCGGAGCTTTCTAATAATGGA	134426
OY	1584	TGATGGTACACTTAATTTCTTGTCGCTCCATCAGATTTCTTAATGATGACTCCAAGT	1643
DB	134427	TGACACCAACAAGCTTTCTTGCTGCTCCATCTCATGATTTCTCCGAGAAAGTACTTAAGT	134486
OY	1644	GACCCGTCACAGACGGCTATATGATGGTTGTTCTCAAGTTA	1683
DB	134487	TCCTGGGAACCTAGCATCTCTGGAGTTATTTCTAGCTTTA	134526
RESULT	10		
LOCUS	ATCHRIV34	196107 bp	DNA linear PLN 16-MAR-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 34.		
ACCESSION	AL161534		
VERSION	AL161534.2	GI:7267939	
KEYWORDS			
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana.		
REFERENCE	Arabiopsis thaliana; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1 (bases 1 to 119463; 141613 to 196107)		
AUTHORS	Peters,S.A., Van Steveren,M., Dirkse,W., Stiekema,W., Mewes,H.W., Lemcke,K. and Meyer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 196107)		
AUTHORS	EU Arabidopsis sequencing project.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAR-2000) MIPS, D-82152 Martinsried, FRG, E-mail: Blochemie, Am Klopferstritz 18a, D-82152 Martinsried, FRG, E-mail: lemc@emlps.biochem.mpg.de,mayer@emlps.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV33 at the 5' end and an overlap with ATCHRIV35 at the 3' end.		
COMMENT	Location/Qualifiers		
FEATURES			
SOURCE	1..196107		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
	4530..23385		
misc_feature	/note="Overlap to BAC T1P17; please refer to this entry for analysis and annotation of the gene fragment at the end of this BAC"		
	4533..23387		
misc_feature	/note="overlap to BAC T4C9; please refer to this entry for analysis and annotation"		
	5637..6595		
gene	/gene="At4g12380"		
	complement(join(5637..5879,6044..6259,6572..6595))		
gene	/gene="At4g12380"		
	complement(join(5637..5879,6044..6259,6572..6595))		
CDS	/note="contains EST gb:T04352"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAB78281.1"		
	/db_xref="GI:7267940"		
	/translation="MTYKRLMLLHKRRKNKOMTEIIRNVPTDEIQEOKNKREER		

LKSALEEGEESDCFEVFGSYDVSYNSGLINISNGRVSRYFVIYKTKFIPIWMNEDFLI  
 KVTWGLGVQVALESYVGLMILCSHQLVTRSPARHOVYISLPRHSARSAPVQVLVVSRL  
 "

exon complement(55637. .5879)  
 /gene="AT4g12380"  
 /number=1  
 Intron complement(5880. .6043)  
 /gene="AT4g12380"  
 /number=1  
 exon complement(6044. .6259)  
 /gene="AT4g12380"  
 /number=2  
 Intron complement(6260. .6571)  
 /gene="AT4g12380"  
 /number=2  
 exon complement(6572. .6595)  
 /gene="AT4g12380"  
 /number=3  
 gene 8679. .9299  
 /gene="AT4g12390"  
 8679. .9299  
 CDS /gene="AT4g12390"  
 /note="similarity to pectinesterase - Citrus sinensis,  
 PID:g2098711  
 Contains Prokaryotic membrane lipoprotein lipid attachment  
 site AA34-44  
 contains EST gb:AI992989.1, R90563, AI100571, T88173,  
 H35998, AA067535"  
 /codon\_start=1  
 /product="putative protein"  
 /protein\_id="CAB78283.1"  
 /db\_xref="GI:7267941"  
 /translation="MEPKLTHCYCLILFLPLICOSTIARPSSPNDSSINFTVSSC  
 RHYRYOTLCYCLAAPADKIRRNNOIACOTPLATVIVOSTTITVYCKITARIKRR  
 EYLAIVDCVENLGGLEMLAOSMRELQVSGSGRDEPLRSLNVEVWSAALDET  
 TCDLGFQKWDGVVKSARIRRVAVARVTSNALALVNRFAARRKS"  
 8679. .9299  
 /gene="AT4g12390"  
 /number=1  
 10992. .13365  
 /gene="AT4g12400"  
 complement(join(10992. .11105,11180. .11306,11384. .11693,  
 11783. .12071,12529. .13365))  
 /gene="AT4g12400"  
 complement(join(10992. .11105,11180. .11306,11384. .11693,  
 11783. .12071,12529. .13365))  
 /gene="AT4g12400"  
 /note="strong similarity to stress-induced protein still  
 -Glycine max, PID:g872116  
 contains EST gb:R64973, T20791, T22244, AA040981,  
 AI100596, T46453, AI100738, T20868, AA395622"  
 /codon\_start=1  
 /product="stress-induced protein still-like protein"  
 /protein\_id="CAB78283.1"  
 /db\_xref="GI:7267942"  
 /translation="MAEEAKSKGNAAFSSGDYATITHTPEAINLSPTNHLIYSNRS  
 SYSLHRYEALSDAKTIELKPDMSKYSRLGAFLGSKDEAVDSQKGLIDPS  
 NEMKSLGLADASBRVSSKSNPYDAPQGRKEMKRLADPESTRYILEDDDDVYKIMPEL  
 ORNPNNILYMKDRVKAKGLVNLVAFGSSGSDTEKKEADEKKEPEPEPELVE  
 EERQKERRERKALKEGEGVAYAKKDKGRAVEHYTKAMELDDDISLTVRAVYLE  
 MKEYEKIECDCAVERGRELISDFKIMARLTPKGSALVMAKSCDFEALTFPOK  
 ALTEHRNPDTLKLNDAEVKYKLEOEEDPTIAEERREKNGEFPKQKXPALVKH  
 SEATKRNPNVRAVSNAACYTKLGALEPKLDAEKICELDSPSTKYSRKGALQFPA  
 KEYDKAMETVOEGIKHPKNOEFLDGRVRCYEOJNKASRGDLTPEELKERAKAMODP  
 EVONILSDPVNRQVLVDFQENPKAOSHNMNPMWNKIQKLVSAGIVQR"  
 complement(10992. .11105)  
 /gene="AT4g12400"  
 /number=1  
 Intron complement(11106. .11179)  
 /gene="AT4g12400"  
 /number=1  
 exon complement(11180. .11306)  
 /gene="AT4g12400"

/number=2  
 complement(11307. .11383)  
 /gene="AT4g12400"  
 /number=2  
 exon complement(11384. .11693)  
 /gene="AT4g12400"  
 /number=3  
 Intron complement(11694. .11782)  
 /gene="AT4g12400"  
 /number=3  
 exon complement(11783. .12071)  
 /gene="AT4g12400"  
 /number=4  
 Intron complement(12072. .12528)  
 /gene="AT4g12400"  
 /number=4  
 exon complement(12529. .13365)  
 /gene="AT4g12400"  
 /number=5  
 tRNA 13550. .13623  
 /product="tRNA-Asn"  
 /note="codon recognized: GUU"  
 15082. .15555  
 /gene="AT4g12410"  
 complement(15082. .15555)  
 /gene="AT4g12410"  
 complement(15082. .15555)  
 /gene="AT4g12410"  
 /note="similarity to auxin-induced protein 10A -Glycine  
 max, PID:g255579  
 contains EST gb:AI93046.1"  
 /codon\_start=1  
 /product="putative protein"  
 /protein\_id="CAB78284.1"  
 /db\_xref="GI:7267943"  
 /translation="MKRVRGFKIGHREYKIKWIPRNRCPTRTINPYGINSARCLS  
 RGARRLCGSSKKNPGONOIRLGDPKSKSNRVYPGHLVAVHGESDDDDRRVVPVYF  
 NHPLEGLLEOAEVHGDPGRITIPCRVSDFEKVLRIAMAMHCRKNSYKII"  
 complement(15082. .15555)  
 /gene="AT4g12410"  
 /number=1  
 22068. .24995  
 /gene="AT4g12420"  
 complement(join(22068. .22167,22248. .22447,22538. .22616,  
 22697. .23062,23144. .23510,23593. .23731,23817. .24087,  
 24195. .24304,24864. .24995))  
 /gene="AT4g12420"  
 complement(join(22068. .22167,22248. .22447,22538. .22616,  
 22697. .23062,23144. .23510,23593. .23731,23817. .24087,  
 24195. .24304,24864. .24995))  
 /gene="AT4g12420"  
 /note="EST GB:Z37701 marks 3'end  
 strong similarity to pollen-specific protein precursor -  
 common tobacco, PIR2:S22495  
 contains EST gb:N38363, R30492"  
 /codon\_start=1  
 /product="putative pollen-specific protein"  
 /protein\_id="CAB78285.1"  
 /db\_xref="GI:7267944"  
 /translation="MDLEKILLFLFEVNIISFCFADPYSFYNFVSYTASPLGPOD  
 VIALNGKPGPTIYNTRENIIVVVRNKLDDGLLHNMGIQORVSMODGVATGNCPI  
 PKNNWYEPQVOKOIGSEFYFSPSLHQRAAGGSPVNPRAIIPVDFSPDDITVY  
 TIGWYIRNTHLAKKALDDGDLGHPGVILNGGPRRYRNTLVADGIDPEPTTVYHG  
 KTVLRVSNVGISSLNRIQGHMLVALESGSTVOONVTSIDIHGOSTSFLVTHD  
 OMASDVIYVASARVNETIWRRTVGVLKVTYSKGAQKQLDPGPDDEDFKFSN  
 QARSIRNVSASGARPNPOGSEFKGSIIVTVVYLNNMPVTIGSKRTTLNGISFN  
 PSTPIRLADKLKVDYVKLDPEKRLNPKAVASIIINGRGMEVYLONDRKMS  
 YHMSGYAFVVGMDYGEETENSRCGYNNMGSIARSTIOVYGAASATILISDNGANN  
 LRTEHDSWYIGQETVYRVVNPDENKTERFHPDNVLTCGALSRLQRPQKVSASAS  
 IGFTSLSNVVALVMMMLQH"  
 complement(22068. .22167)  
 /gene="AT4g12420"  
 /number=1  
 exon







```

* 25801 25900: gap of unknown length
* 25901 27016: contig of 1116 bp in length
* 27017 27116: gap of unknown length
* 27117 28154: contig of 1038 bp in length
* 28155 28254: gap of unknown length
* 28255 29648: contig of 1394 bp in length
* 29649 29749: gap of unknown length
* 29749 31439: contig of 1691 bp in length
* 31440 31539: gap of unknown length
* 31540 32736: contig of 1197 bp in length
* 32737 32836: gap of unknown length
* 32837 34705: contig of 1869 bp in length
* 34706 34805: gap of unknown length
* 34806 36369: contig of 1564 bp in length
* 36370 36469: gap of unknown length
* 36470 37820: contig of 1351 bp in length
* 37821 37920: gap of unknown length
* 37921 39085: contig of 1145 bp in length
* 39086 39165: gap of unknown length
* 39166 40330: contig of 1165 bp in length
* 40331 42100: contig of 1670 bp in length
* 42101 42200: gap of unknown length
* 42201 43481: contig of 1281 bp in length
* 43482 43581: gap of unknown length
* 43582 45285: contig of 1704 bp in length
* 45286 45385: gap of unknown length
* 45386 47189: contig of 1804 bp in length
* 47190 47289: gap of unknown length
* 47290 48769: contig of 1480 bp in length
* 48770 48869: gap of unknown length
* 48870 50612: contig of 1743 bp in length
* 50613 53217: gap of unknown length
* 53218 53317: contig of 2505 bp in length
* 53318 54410: gap of unknown length
* 54411 54510: contig of 1093 bp in length
* 54511 55559: gap of unknown length
* 55560 55659: gap of unknown length
* 55660 56993: contig of 1334 bp in length
* 56994 57093: gap of unknown length
* 57094 58794: contig of 1701 bp in length
* 58795 58894: gap of unknown length
* 58895 60396: contig of 1504 bp in length
* 60399 60498: gap of unknown length
* 60499 61727: contig of 1229 bp in length
* 61728 61827: gap of unknown length
* 61828 63020: contig of 1193 bp in length
* 63021 63120: gap of unknown length
* 63121 64340: contig of 1420 bp in length
* 64341 64640: gap of unknown length
* 64641 65892: contig of 1252 bp in length
* 65893 65992: gap of unknown length
* 65993 68071: contig of 2079 bp in length
* 68072 68171: gap of unknown length
* 68172 69557: contig of 1786 bp in length
* 69558 70057: gap of unknown length
* 70058 72094: contig of 2037 bp in length
* 72095 72194: gap of unknown length
* 72195 73883: contig of 1689 bp in length
* 73884 73983: gap of unknown length
* 73984 75384: contig of 1401 bp in length
* 75385 75484: gap of unknown length
* 75485 76673: contig of 1189 bp in length
* 76674 78925: gap of unknown length
* 78926 79025: contig of 2152 bp in length
* 79026 80869: gap of unknown length
* 80870 80969: gap of unknown length
* 80970 82475: contig of 1506 bp in length
* 82476 82575: gap of unknown length
* 82576 83646: contig of 1071 bp in length

```

```

Query Match 1.7%; Score 45.4; DB 2; Length 161477;
Best Local Similarity 50.6%; Pred. No. 0.2;
Matches 137; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

QY 158 GAGGCTTGGGTTTATATACAGACATGTGGAGAGAGTCGACGCCAAGAGGCTC 217
    |||||
Db 155197 GTGGCTTGGCTTCTTCAATTCAGAACTGGACATGACACAGGAGCCAGGTGCTA 155256
    |||||

QY 218 TTCAAGAGCAAAATTTAAATGAAGTCAATTAAGATGCAATTCAGACGACGCAAAAC 277
    |||||
Db 155257 TGCTGTGCGCGGGTGATGGCAAAACCTTAATAAGATAGGCTACG---GCAAGGCTTAACC 155313

QY 278 CTTCCTAAGAGCTATGAGTGAGTGAGTAATGCGCCCTTAATGCTCAAGATGAGCTGAGG 337
    |||||
Db 155314 CTACACACCCCTCTGTGGGTGGGTGCTTTGAGACCTTAACACTGCTGCGCCCTGAGCA 155373
    |||||

QY 338 AAGAGTTCAGCAAGTTTGGAAAATCGAGATTTTAACTTTCAGAGAAAGCAAGACAG 397
    |||||
Db 155374 GAGAAATTCATCGCTTGGGAGATTCGACCAATCGATCAAGCAAGAGACAGCTTTG 155433
    |||||

QY 398 CTTCATGATGATTAATGATGATGATGC 428
    |||||
Db 155434 CTACATCCAGTACGAGAGGCTGTGATGCAC 155464
    |||||

```

```

RESULT 14
AX083744
LOCUS AX083744 1141 bp DNA linear PAT 28-FEB-2001
DEFINITION Sequence 22 from Patent WO0111061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequences.
REFERENCE
1 (bases 1 to 1141)
AUTHORS
Kunst, L. and Clemens, S.
TITLE
Regulation of embryonic transcription in plants
JOURNAL
Patent: WO 011061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES
Source
1. 1141
/organism="synthetic construct"
/db_xref="taxon:32630"
1. 1141
/promoter="consensus sequence of A.t., L.a., and B.n. FAEL
promoters"
BASE COUNT 123 a 32 c 42 g 112 t 832 others
ORIGIN
Query Match 1.7%; Score 45.2; DB 6; Length 1141;
Best Local Similarity 10.6%; Pred. No. 0.21;
Matches 82; Conservative 276; Mismatches 413; Indels 0; Gaps 0;

QY 379 CTCGAGAAACGCAAGAGCTTCATGATTAATGAGATGATGATGCTTTACAGGCT 438
    |||||
Db 54 SKRSKWTWAMRYCKYRRWYNNKSRMWKGMKKKWKYBCANNYSBRHARBMKDKTAYBMT 113
    |||||

QY 439 AAGAGCATGAATGAAGAGCTATGGGTGAGCTTTTGGCGTGTATTTCTCCGGTCA 498
    |||||
Db 114 MTRKMGTRRHRYWYRAMBDTVDDHYVTAMNNATTCMDKDRTRRMWKKNNNT 173
    |||||

QY 499 CAAGCCCAAAAAAACAACAATGGGCTGCTTACGATTAACAAGAAATGCAATATGAAT 558
    |||||
Db 174 GMDDDTKRYHMMNNNGCBYTYWAVRYTDDMSKRNRYGMBWKNNSYVTTYWYWDGM 233
    |||||

QY 559 CATAAACCGACATATCTCATCATATGAAGATTTAAAGAGATGCTCCAGCAAGTAGA 618
    |||||
Db 234 CKRKVBRWVTRGRMKNRYWAVABTAHRRRYNNGWBTAAAYRRWTNNNNNAKAMCKRAK 293
    |||||

QY 619 GTTCTGTGATTTGGTTCCTCTCTACTGCTACACAATGATGAGCAAAATTTCTGCAC 678
    |||||
Db 294 YMGWNRABVNSTCTTKRSKTYRTVRSWMANNCBAGDANKDHRWKKWSAAGYWNNNNN 353
    |||||

```



[illegible]

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 00:45:48 ; Search time 3748 Seconds  
(without alignments)  
11692.899 Million cell updates/sec

Title: US-09-920-705-2  
Perfect score: 2706  
Sequence: 1 atggcgtactatgaagcc.....cgcgtgacagggccttga 2706

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estbda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	403.6	14.9	701 10 BE037904	BE037904 AA06B09 A
2	403	14.9	767 17 BH552322	BH552322 BOGH093TF
3	350.2	12.9	635 9 AU237346	AU237346 AU237346
4	334.6	12.4	724 17 BH552335	BH552335 BOGH093TR
5	330.8	12.2	602 12 BG543956	BG543956 E1706 Ch1
6	323.4	12.0	416 17 BH851246	BH851246 SALK_0727

7	303	11.2	582	10	AV829561
8	194	7.2	561	12	BF266387
9	177.8	6.6	653	10	AM96401
10	176	6.5	630	12	BG447636
11	170.8	6.3	647	14	BO696544
12	170.8	6.3	662	12	BF641487
13	170.2	6.3	685	13	BI954416
14	165.6	6.1	617	12	BG649121
15	161	5.9	491	10	AV548323
16	154.2	5.7	381	14	BQ697581
17	145.4	5.4	606	10	BE594381
18	139.8	5.2	493	12	BF639297
19	133.4	4.9	472	14	N96947
20	129.8	4.8	691	9	AL505369
21	123.2	4.6	673	12	BG600897
22	121.8	4.5	429	10	AV406694
23	118.6	4.4	751	12	BG591991
24	118.4	4.4	317	14	C24205
25	112.8	4.2	430	12	BF585815
26	112	4.1	475	9	AU228388
27	108	4.0	526	10	AM033033
28	106.6	3.9	423	10	AV407805
29	100.6	3.7	589	12	BG649379
30	96.4	3.6	644	12	BE278330
31	95.4	3.5	670	14	BQ401768
32	95	3.5	586	10	BE499504
33	91.2	3.4	820	17	BH715029
34	85.4	3.2	539	10	AM596129
35	84.6	3.1	457	9	AL378093
36	84.6	3.1	473	9	AL378092
37	80.8	3.0	439	14	BO489608
38	76.6	2.8	480	10	BA834854
39	73.2	2.7	334	13	BI751900
40	72	2.7	577	12	BF003832
41	71.8	2.7	493	10	AV529289
42	71.8	2.7	528	9	AU237648
43	69.4	2.6	430	17	BH251458
44	65.4	2.4	475	9	AU174330
45	62.8	2.3	769	17	BH442944

## ALIGNMENTS

RESULT 1  
BE037904  
LOCUS BE037904 701 bp mRNA linear EST 07-JUN-2000  
DEFINITION AA Arabidopsis thaliana cDNA 5', mRNA sequence.  
ACCESSION BE037904  
VERSION BE037904.1 GI:8332920  
KEYWORDS EST.

## ORGANISM

thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

Bohner, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira, H., Kawasaki, S., McColough, A., Michalowski, C.B., Palacio, C., Scarra, G., Wheeler, M. and Zepeda, G.R.  
Functional Genomics of Plant Stress Tolerance  
Unpublished (2000)

## JOURNAL

COMMENT  
Contact: Michalowski, C.B.  
University of Arizona  
Bio Sciences West room 513, Tucson, AZ 85721, USA  
Tel: 520-621-7982  
Fax: 520-621-1697  
Email: chmdu.arizona.edu

FEATURES  
Source  
1. 701  
Best Blastx match: 'gb|AA064314.1 (AC002335) hypothetical protein [Arabidopsis thal.]. 165 3e-40'. An open reading frame exists.  
Insert length: 2 Std Error: 0.00.  
Location/Qualifiers

Email: [ecolomnec@u.wisc.edu](mailto:ecolomnec@u.wisc.edu)  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF

Db 40 AACATGCCCTCAGATGATC 22



LOCUS	DEFINITION	Accession	Version	Keywords	Organism	Reference	Authors	Title	Journal	Comment
RESULT 3 AU237346										
LOCUS	AU237346	635 bp	mRNA	linear	EST 01-APR-2002					
DEFINITION	AU237346 RafL16 Arabidopsis thaliana cDNA clone RafL16-16-A15 5'									
Accession	AU237346									
Version	AU237346.1	GI:19876515								
Keywords	EST									
Organism	thale cress. Arabidopsis thaliana									
Reference	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 635)									
Authors	Seki, M., Natusaka, M., Ishida, J., Kamiya, A., Setou, M., Nakajima, M., Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carlinici, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.									
Title	Large scale analysis of Arabidopsis full-length cDNA									
Journal	Unpublished (2002)									
Comment	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: mseki@rtc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FIC-1 vector (Carlinici et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site ( <a href="http://www.gsc.riken.go.jp/e/plant/index_e.html">http://www.gsc.riken.go.jp/e/plant/index_e.html</a> ) for further details.									
FEATURES										
source	1. 635 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="RafL16-16-A15" /clone_1ib="RafL16" /lab_host="DH10B" /note="Site_1: BamHI; Site_2: SalI; dark-grown"									
BASE COUNT	172 a 137 c 144 g 179 t 3 others									
ORIGIN										
Query Match	12.9% Score 350.2; DB 9; Length 635;									
Best Local Similarity	97.9% Pident. NO. 1.2e-97;									
Matches 374; Conservative	0; Mismatches 6; Indels 2; Gaps 2;									
1	ATGGCGTTATCTATGAAGCCATTGCAGAGCCGATGATTCGGTTCACATCAACAATTT	60								
255	ATGGCGTTATCTATGAAGCCATTGCAGAGCCGATGATTCGGTTCACATCAACAATTT	314								
61	TGGGTGCGTACCCCTAACGCCGAGAGACGACAGATCAGATCTGACCGAATTTGGAGA	120								
315	TGGGTGCGTACCCCTAACGCCGAGAGACGACAGATCAGATCTGACCGAATTTGGAGA	374								
121	TACGGCATATGTATGAATCAGCGTGTATTTCTTCACAGAGCTTTGGCTTATATATAC	180								
375	TACGGCATATGTATGAATCAGCGTGTATTTCTTCACAGAGCTTTGGCTTATATATAC	434								
181	AGACATGTGAGAGAACGAGTGCAGACCAAGAGGCTCTTCAAGAGCAAAATTGAATGA	240								
435	AGACATGTGAGAGAACGAGTGCAGACCAAGAGGCTCTTCAAGAGCAAAATTGAATGA	494								
241	AGTCAAAATTGAATGCAATACGACGACGACGCAAAACCTTTGAAGAGTCTAT -GGTGG	299								
495	AGTCAAAATTGAATGCAATACGACGACGACGCAAAACCTTTGAAGAGTCTAT -GGTGG	554								
300	TGGAATGGCCCTTAATGCTCTCCAGAGATGACCTTGAGAGAAAGTTCAAGACTTTGGAA	359								
555	TGGAATGGCCCTTAATGCTCTCCAGAGATGACCTTGAGAGAAAGTTCAAGACTTT -GGAA	613								

[illegible]

QY	903	AGATATGGTCATTATGAAATCCCTTAAAGTATGTAATCTTCTCTCTCAATCTTACTGAAATCTTCGG	962
Db	529	AGGTATGTTTCACTACACTGATCTCTTCAAT	554
QY	963	GTCTATGAGGCCCCCTCAGAGGTAGCAATGAGCGTTTCATATATGTTGTCAGATACATGA	1022
Db	555	-----TGAAGCATTCACAGAGGTAGCAATGACGCGTTCTTATATATGTTTCAGATTAACATGA	609
QY	1023	CGTTGTTGGTAGGAGCCCAATCTGAGAGGCGCATCTGCAATATGTAATGTAATCTCTCC	1082
Db	610	CGTTGTTGGTAGGAGCCCAATCTGAGAGGCGCGTTCTCAATATGGAATGGATGGATATCTCC	669
QY	1083	ATCTCCACAGAGCACTGGATCTCTCCATCTCTCTGTCACAGAGGTAGGAGCGCCCT	1137
Db	670	ATCTCCACAGAGCACTGGATCTCTCCATCTCTCTGTCACAGAGGTAGGAGCAACCT	724
RESULT 5	BSG43956	602 bp	linear
DEFINITION	BSG43956	602 bp	linear
LOCUS	E1706	Chinese cabbage etiolated seedling library	EST 01-MAY-2002
ACCESSION	subsp. pekinensis	cDNA clone E1706, mRNA sequence.	
VERSION	BSG43956		
KEYWORDS	BSG43956.1	GI:20374936	
SOURCE	EST		
ORGANISM	Brassica rapa subsp. pekinensis.		
	Brassica rapa subsp. pekinensis		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
REFERENCE	1 (bases 1 to 602)		
AUTHORS	Ryu,S.H., Kang,Y.S., Kang,C.-h., Kim,C.Y., Choi,Y.J., Lee,S.-H.,		
	Baek,U.D., Lee,S.Y., Cho,M.J. and Lim,C.O.		
TITLE	Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNA		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Lim, C.O.		
	Plant Molecular Biology & Biotechnology Research Centre		
	Gyeongsang National University		
	#900 Gaeje-dong, Jinju 660-701, Korea		
	Tel: 82 55 751 6255		
	Fax: 82 55 759 9363		
	Email: colim@nongae.gsnu.ac.kr		
	Seq primer: T7.		
FEATURES	Location/Qualifiers		
source	1..602		
	/organism="Brassica rapa subsp. pekinensis"		
	/cultivar="Jangwon"		
	/db_xref="taxon:51351"		
	/clone="E1706"		
	/clone_lib="Chinese cabbage etiolated seedling library"		
	/issue_type="Etiolated seedling"		
	/lab_host="Xf-1 Blue"		
	/note="Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"		
BASE COUNT	147 a 144 c 166 g 145 t		
ORIGIN			
Query Match	12.2%;	Score 330.8;	DB 12;
Best Local Similarity	75.6%;	Pred. No. 1.3e-91;	Length 602;
Matches	440;	Conservative 0;	Mismatches 132;
		Indels 10;	Gaps 2;
QY	67	GGTAGCTACGCGCGGAGACGACGAGCTGACAGTGTGTTGGAAGTACGCG	126
Db	13	GGCAGTCTGACCCCGGACTACGATCGATCGATCTGCCGATGTTCGGAGGTTGGCG	72
QY	127	GATATTGTTAATACAGGCTGTATTCTTCACAGAGCTTTGGTTTATATCTACAGACAT	186
Db	73	GATATCGACAGCGCTACCGCTATTCTTCCCGTGGCTTCGATTAATCTTCAACGAGCG	132
QY	187	GTGAGGAGAGAGTCCGACGACCAAGAGGCTTTCAGAGCAAAATTTGAATGAGACAA	246
Db	133	GTTGAGGAGGCGGTAGCGCGGAGAGAGAGCTTTAGGTTGCGCAATCTAAACGGAGTCC	192
QY	247	ATTAAAGTCCAAATTCGACGACGCGGCAAAACCTTTGTAAGAGTCTAATGAGTGGGTGAAATC	306

	BH851246	RESULT 6 BH851246
	LOCUS	
	DEFINITION	SALK_072730..56.00.x Arabidopsis thaliana TDNA insertion lines
	ACCESSION	BH851246
	VERSION	GSS.
	KEYWORDS	
	SOURCE	
	ORGANISM	thale cress. <i>Arabidopsis thaliana</i> Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes I; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 416) Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadgilab, .C., Jeske,A.A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R. A sequence-indexed library of Insertion Mutations In the Arabidopsis genome Unpublished (2001)
	JOURNAL COMMENT	Contact: Joseph R. Ecker Saik Institute Genomic Analysis Laboratory (SIGNAL) The Saik Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel.: 858 453 4100 x1752 Fax: 858 558 6379 Email: eckeresalk.edu This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of AtZg49410. Class: TDNA tagged.
FEATURES		Location/Qualifiers
source		1..416 /organism="Arabidopsis thaliana" /strain="Columbia 0" /db_xref="taxon:3702" /clone="-SALK_072730..56.00.x" /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.saik.edu/tdna_protocols.html"
COUNT		136 a 109 c 78 g 93 t

## ORIGIN

Query Match 12.0%; Score 323.4; DB 17; Length 416;  
Best Local Similarity 99.7%; Pred. No. 2.1e-89;  
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2382 GCCAACCAAACTGTGTCCTCAATTCAGCAAGTATATTCATGGCGTAAC 2441  
DB 1 GCCAACCAAACTGTGTCCTCAATTCAGCAAGTATATTCATGGCGTAAC 60  
QY 2442 AACCAATCAGCAGCACTTAACCTTCCCAATTTCAAGTGCATGCACACACAGC 2501  
DB 61 AACCAATCAGCAGCACTTAACCTTCCCAATTTCAAGTGCATGCACACACAGC 120  
QY 2502 AGATAAGCAATTTAGAGCCCAAAACAGCACTAGCATTCGAGCCTATGATCTTGG 2561  
DB 121 AGATAAGCAATTTAGAGCCCAAAACAGCACTAGCATTCGAGCCTATGATCTTGG 180  
QY 2562 GGATGTGAGGTTACAGAGAGGAGTGATTAAGATCAAGATACCAATCAACT 2621  
DB 181 GGATGTGAGGTTACAGAGAGGAGTGATTAAGATCAAGATACCAACT 240  
QY 2622 ACAATTTGAGCAAACTTCTCTCAGATACAGCAAGAAACAGCACTCTTCAG 2681  
DB 241 ACAATTTGAGCAAACTTCTCTCAGATACAGCAAGAAACAGCACTCTTCAG 300  
QY 2682 TACTCCGCTGGACAGGGGCTTGA 2706  
DB 301 TACTCCGCTGGACAGGGGCTTGA 325

## RESULT 7

AV829561 582 bp mRNA linear EST 01-APR-2002  
LOCUS AV829561 RAF19 Arabidopsis thaliana cDNA clone RAF109-47-P10 5',  
DEFINITION mRNA sequence.

ACCESSION AV829561  
VERSION AV829561.1 GI:19871621  
KEYWORDS EST.

## SOURCE

## ORGANISM

thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 582)

## REFERENCE

## AUTHORS

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Setou, M., Nakajima, M.,  
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.  
and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length cDNA (2002b)  
Unpublished (2002)

## JOURNAL

## COMMENT

Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@tc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
and XhoI was ligated to modified Lambda PhC-1 vector (Carninci et  
al., submitted for publication) digested with BamHI and SalI. This  
clone is in a modified Bluescript vector. Please visit our web  
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further  
details.

## FEATURES

## source

1..582  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/clone="RAF109-47-P10"  
/clone\_11b="RAF19"  
/dev\_stage="plants at various developmental stages from  
germination to mature seeds"

## ORIGIN

Query Match 11.2%; Score 303; DB 10; Length 582;  
Best Local Similarity 98.8%; Pred. No. 6.8e-83;  
Matches 325; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 ATGGCGTTATCTATGAGCAAGCATTCAGAGCCGAGTTCGGTTCCAGTCAACATCT 60  
DB 254 ATGGCGTTATCTATGAGCAAGCATTCAGAGCCGAGTTCGGTTCCAGTCAACATCT 313  
QY 61 TGGGTCGTAGCCTTAACGCGGAGAGACAGACAGTCAAGTCAAGCTTTTGGAGA 120  
DB 314 TGGGTCGTAGCCTTAACGCGGAGAGACAGACAGTCAAGTCAAGCTTTTGGAGA 373  
QY 121 TACGGCGATTTGATGATACAGGCTGTCTTTCACGAGGCTTTCGTTTATATCTAC 180  
DB 374 TACGGCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433  
QY 181 AGCAGTGTGAGAGAGACAGTTCGAGCCAAAGAGGCTTCAAGAGCAATTTGAATGA 240  
DB 434 AGCAGTGTGAGAGAGACAGTTCGAGCCAAAGAGGCTTTCAGAGCAATTTGAATGA 493  
QY 241 AGTCAAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298  
DB 494 AGTCAAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553  
QY 299 GTGGAATCGGCCCTATCTCTCCAGGAT 327  
DB 554 GTGGAATCGGCCCTATCTCTCCAGGAT 582

## RESULT 8

BF266387 561 bp mRNA linear EST 23-OCT-2001  
LOCUS BF266387  
DEFINITION HV\_CEA0015A03f Hordeum vulgare seedling green leaf EST library  
HVCNDA0004 (Blumeria challenged) Hordeum vulgare cDNA clone  
HV\_CEA0015A03f, mRNA sequence.

ACCESSION BF266387  
VERSION BF266387.3 GI:16334924  
KEYWORDS EST.

## SOURCE

## ORGANISM

Hordeum vulgare.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
; Triticeae; Hordeum.  
1 (bases 1 to 561)

## REFERENCE

## AUTHORS

Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wel, F., Begum, D.,  
Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi,  
D.W., Fenton, R.D., Oates, R., and Main, D.  
Development of a genetically and physically anchored EST resource  
for barley genomics: Blumeria infected incompatible (Ma13)  
seedling leaf cDNA library  
Unpublished (2001)

## JOURNAL

## COMMENT

On Nov 17, 2000 this sequence version replaced g1:13262704.  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Total hg bases = 389

Seq primer: AATTACCTCTACTAAGGC  
High quality sequence stop: 476.  
Location/Qualifiers  
1..561

## FEATURES

## source

/organism="Hordeum vulgare"  
/cultivar="CI16155 (Ma13)"



OY	1235	CTCCATAGGCGTTCGCATGACAGGTCAATTGGTGACAGTTCAAGTTGCTCCTAGACACTTAACC	1294
Db	259	ATCCTTAGCGAGGAGGCGGTATCCCTTGCTTCAGACAACAGAGGATTACAGTTCGAGTAC	318
OY	1295	GTCGCCCCCCCTGAATTCCTGATACACATATGAGAGGAAATGATTTGCCAAGGCTGGAACTCCCG	1354
Db	319	ATGCTGTGC---AACCTGATCATATATATGCGCGGACTATTATTCGCAAAGAGAACTCTCG	375
OY	1355	TCTGTTTGTGCTGCTTTGTGTACCTATGCGGAAAAGGAGATTTGAAACATAACTGCTGAGGTGC	1414
Db	376	TTTTTGCTGCTAGATGATCCCTCTTAGGAAAGGATCGGGACAGACACTTCCTAGACTTG	435
OY	1415	TCAATTGTTACGACAAGAATCGAATTTGAATATGCTCGCTAAACATTACCGCTGGCCATTG	1474
Db	436	TGATTTCTCAGCTAGAGACGGGATTTGGATTACACTCGACAGACACTATGCTGATCCAATTG	495
OY	1475	GAGTGTGAGATCGTTTTTTTTCGTACCCAGACAGGGAAGAAATTTTGGCTCTTACACTCAAT	1534
Db	496	ACTTTGAGATTTGTTTTTTCTTCGCTATAGTGAAAAATTTTGGTTCACTACCTGAAT	555
OY	1535	TTCCCGCGTACCTTAGCT-CAAAAGATCGGCGGCGGTGTCGCAAAATTAGATGAGTGTACA	1593
Db	556	TTCCAGCCTACCTTTGGTTCGCAAAAATCGTGTCTGTGTTGCAAAATTTTGAGATAC-A	614
OY	1594	ACTTTATTCCTTGTGTC	1609
Db	615	ACTTTATTCCTTGTGTC	630
RESULT 11	B0696544	647 bp	mRNA linear EST 15-JUL-2002
LOCUS	B0696544		
DEFINITION	NXPV_042_E01.F NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda		
VERSION	B0696544		
KEYWORDS	B0696544.1 GI:21821860		
SOURCE ORGANISM	EST. loblolly pine. Pinus taeda		
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.		
AUTHORS	Sederoff,R.		
JOURNAL COMMENT	Molecular Basis of Wood Formation In the Pine Megagenome Unpublished (2000) Contact: Johnson, Arthur North Carolina State University Tel.: 919 515 7800 Fax: 919 515 7801 Email: ajohnson@unity.ncsu.edu Seq primer: T3		
FEATURES	Location/Qualifiers		
Source	1..647 /organism="Pinus taeda" /strain="Coastal plain loblolly pine from North Carolina" /db_xref="taxon:3352" /clone="NXPV_042_E01" /clone_id="NXPV (Nsf Xylem Planings wood Vertical)" /tissue_type="Xylem" /cell_type="Planings (secondary)" /dev_stage="transitional" /lab_host="XL1-Blue" /note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI ; The library is from early (spring) secondary wood, taken from a ten year old tree in the transitional phase. The tree is a kind gift of the Westvaco Corporation. Secondary xylem was harvested from the tree by peeling back the bark and primary xylem and then removing the underlying tissue with a block plane. NOTE: The sequences contain a 'cdna adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGGACGAG'."		
BASE COUNT	177 a	137 c	136 g 172 t 25 others

BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches
175 a	153 c	137 g	194 t	3 others
<p>expanded <i>M. truncatula</i> leaves of plants fed upon by Spodoptera exigua (beet armyworm) for 24 hours. Systemic (undamaged leaves from injured plants) and wounded leaves were harvested and pooled.</p>				
		6.3%;	Score 170.8; DB 12;	Length 662;
		Best Local Similarity 62.9%;	Pred No.2.1e-41;	
		Matches 295; Conservative	0; Mismatches 170;	Indels 4; Gaps 2

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF064H03.IN"
/clone_1bp="insect herbivory"
/tissue_type="local and systemic leaves"
/dev_stage="mature"
/note="Vector: Lambda Zap. Library was produced from fully

```

1 (bases 1 to 685)  
Wing, R., Close, T. J., Kleinbotts, A., Wise, R., Chin, A., Begum, D.,  
Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons,  
J., Oates, R. and Main, D.  
Development of a genetically and physically anchored EST resource  
for barley genomics: Blumeria infected Morex (compatible) seedling  
cDNA library  
unpublished (2001)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
email: rwing@clemson.edu





```

Db      182  AAATT---TGAGATACAACTTATCTGTGCTGCTCCAGATTCCGACGATGTT 238
QY      1636  CTCCAAGTGAACCCCTCAAGAACGGCTATATGTTTCTTCAAGTTACCCCGCCAGCC 1695
Db      239  CTGAAGTCAACCGGACCGAAGCTATATGTTGTTCTCAAGTTGACCCAGTACAA 298
QY      1696  GTTCCCTGTACAGCATATACAGACAGAAATCTCAGTCAATCCTTCATTTATATGAT 1755
Db      299  AGTGTGTGCGCCGCGACCAATCATCATTTGCTGTACCGTCAATCATATATATGAC 358
QY      1756  CAAGCCCGGAGATTCACCTGC 1775
Db      359  CAGATGCCTCTTCACAGGC 378

```

## RESULT 15

AV548323/c

LOCUS AV548323 491 bp mRNA linear EST 06-SEP-2000

DEFINITION AV548323 Arabidopsis thaliana roots Columbia Arabidopsis thaliana

ACCESSION AV548323

VERSION AV548323.1 GI:8719736

KEYWORDS EST.

SOURCE

thale cress.

Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 491) Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

A large scale analysis of cDNA in Arabidopsis thaliana: Generation

of 12,028 non-redundant expressed sequence tags from normalized and

size-selected cDNA libraries

DNA Res. 7, 175-180 (2000)

20363093

JOURNAL Contact: Erika Asamizu

MEDLINE The First Laboratory for Plant Gene Research

COMMENT Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES location/Qualifiers

source 1..491

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db\_xref="taxon:3702"

/clone="RZL51908F"

/clone\_lib="Arabidopsis thaliana roots Columbia"

/tissue\_type="roots"

/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 150 a 92 c 96 g 153 t

ORIGIN

Query Match 5.9%; Score 161; DB 10; Length 491;

Best Local Similarity 100.0%; Pred. No. 1.9e-38;

Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2546 AGCCTATGATCTCTGGGATGTGTCAGGATACACAGATGGGAGTGCATTAAGATCAGA 2605

Db 491 AGCCTATGATCTCTGGGATGTGTCAGGATACACAGATGGGAGTGCATTAAGATCAGA 432

QY 2606 GATACAGTCAACACTACATTTTGCAGCAAAACCTTCTTCAGATACAGCAAAACAGC 2665

Db 431 GATACAGTCAACACTACATTTTGCAGCAAAACCTTCTTCAGATACAGCAAAACAGC 372

QY 2666 AGCAACAGTCTAGTACTCCGGCTGCAGAGGGGCTTGA 2706

Db 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGGGCTTGA 331

Search completed: June 19, 2003, 04:25:36

Job time : 3767 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 03:22:45 : Search time 914 Seconds  
(without alignments)  
4344.484 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706  
Sequence: 1 atgcgcgtatctatgaagcc.....cgcgtgacagggccttga 2706

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCr\_NEM\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCrUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEM\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEM\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2706	100.0	2706	US-09-920-705-2	Sequence 2, Appli
2	1896.8	70.1	4593	US-09-920-705-1	Sequence 1, Appli
3	797.8	29.5	801	US-09-920-705-4	Sequence 4, Appli
4	271.4	10.0	349	US-09-920-705-5	Sequence 5, Appli
5	271.4	10.0	3715	US-09-920-705-6	Sequence 6, Appli
6	113.2	4.2	438	US-09-770-444-972	Sequence 972, App
7	42.8	1.6	256	US-09-815-343-1381	Sequence 181, App
8	41.6	1.5	2142	US-09-917-800A-1567	Sequence 1567, Ap
9	39.6	1.5	653	US-10-123-155-438	Sequence 438, App
10	38.6	1.4	1344	US-10-062-254-287	Sequence 287, App
11	38.2	1.4	1049	US-10-123-155-358	Sequence 358, App
12	37.2	1.4	424	US-09-878-574-4075	Sequence 4075, Ap
13	37.2	1.4	1537	US-10-062-254-275	Sequence 275, App
14	37	1.4	1846	US-10-197-666A-77	Sequence 77, Appl
15	37	1.4	1964	US-10-197-666A-79	Sequence 79, Appl
16	37	1.4	2037	US-10-071-766-104	Sequence 104, App
17	36.4	1.3	396	US-09-803-719-708	Sequence 708, App
18	36.4	1.3	406	US-09-803-719-783	Sequence 783, App
19	36.4	1.3	407	US-09-803-719-781	Sequence 781, App

20	36.4	1.3	466	US-09-918-995-11693	Sequence 11693, A
21	36	1.3	1024	US-10-123-155-198	Sequence 198, App
22	36	1.3	2000	US-09-938-842A-2746	Sequence 2746, Ap
23	35.8	1.3	354	US-10-102-534-1033	Sequence 1033, Ap
24	35.2	1.3	637	US-09-778-320-189	Sequence 189, App
25	35.2	1.3	637	US-09-910-689-189	Sequence 189, App
26	35.2	1.3	637	US-10-010-742-189	Sequence 189, App
27	35.2	1.3	1583	US-10-062-254-273	Sequence 273, App
28	35	1.3	1799	US-10-062-254-247	Sequence 247, App
29	35	1.3	1806	US-10-062-254-249	Sequence 249, App
30	34.8	1.3	487	US-10-123-155-528	Sequence 528, App
31	34.8	1.3	1184	US-10-123-155-412	Sequence 412, App
32	34.8	1.3	1728	US-09-794-589-3	Sequence 3, Appli
33	34.4	1.3	734	US-10-184-644-458	Sequence 458, App
34	34.4	1.3	734	US-10-184-644-458	Sequence 458, App
35	34.4	1.3	2552	US-09-799-777-141	Sequence 141, App
36	33.8	1.2	407	US-09-918-995-36239	Sequence 36239, A
37	33.8	1.2	600	US-09-917-800A-857	Sequence 857, App
38	33.8	1.2	867	US-10-123-155-20	Sequence 20, Appl
39	33.6	1.2	1160	US-10-123-155-234	Sequence 234, App
40	33.6	1.2	1757	US-10-062-254-269	Sequence 269, App
41	33.4	1.2	421	US-09-864-761-265	Sequence 265, App
42	33.4	1.2	470	US-09-864-761-867	Sequence 867, App
43	33.4	1.2	479	US-09-864-761-794	Sequence 794, App
44	33.4	1.2	581	US-09-864-761-9417	Sequence 9417, Ap
45	33.4	1.2	2116	US-10-002-600-81	Sequence 81, Appl

#### ALIGNMENTS

RESULT 1  
US-09-920-705-2

Sequence 2, Application US/09920705  
Publication No. US20030079252A1

GENERAL INFORMATION:

APPLICANT: Amasino, Richard M.

APPLICANT: Schomburg, Fritz M.

APPLICANT: Michaels, Scott D.

APPLICANT: Patton, David

TITLE OF INVENTION: Floral Induction Gene

FILE REFERENCE: 960296.97214

CURRENT APPLICATION NUMBER: US/09/920,705

CURRENT FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 2706

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2706)

NAME/KEY: misc\_binding

LOCATION: (46)..(279)

OTHER INFORMATION: RNA Binding Region

NAME/KEY: misc\_binding

LOCATION: (283)..(522)

OTHER INFORMATION: RNA Binding Region

NAME/KEY: misc\_binding

LOCATION: (610)..(852)

OTHER INFORMATION: RNA Binding Region

US-09-920-705-2

Query Match

Best Local Similarity 100.0%; Score 2706; DB 9; Length 2706;

Matches 2706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGGTATCTATTAAGCCATTCAGACCCATGATTCGGCTTCCATCAACAATCTT 60  
DB 1 ATGCGGTATCTATTAAGCCATTCAGACCCATGATTCGGCTTCCATCAACAATCTT 60  
QY 61 TGGCGGTAGCCTTAACCCGAGACAGACAGATCAGATCTGACCGACTTGTGAGACA 120



```
OY 2281 CCAATATGCTCATTAATTAATTCATGACACTGAGGTTCGTAATCAATCCTGTT 2340
DB 2281 CCAATATGCTCATTAATTAATTCATGACACTGAGGTTCGTAATCAATCCTGTT 2340
OY 2341 TCTGAGCCATGTCGACGACATACCAACAGAACGTCGATGCAACCAACCAATATGTT 2400
DB 2341 TCTGAGCCATGTCGACGACATACCAACAGAACGTCGATGCAACCAACCAATATGTT 2400
OY 2401 CCAATTCAGATATGACGACAAAGCTAATTTTCATGCGGTACACAAATCAGGACAGAAC 2460
DB 2401 CCAATTCAGATATGACGACAAAGCTAATTTTCATGCGGTACACAAATCAGGACAGAAC 2460
OY 2461 TTAACCCCTCCCATTTTCAAGCTGCGATGACACACGACGATGATAGGCAAAATTTAGAG 2520
DB 2461 TTAACCCCTCCCATTTTCAAGCTGCGATGACACACGACGATGATAGGCAAAATTTAGAG 2520
OY 2521 CCACAAAACCAAGCAGCTACGATTCGACGCTATGATCTCTGGGATGTCAGGATACACA 2580
DB 2521 CCACAAAACCAAGCAGCTACGATTCGACGCTATGATCTCTGGGATGTCAGGATACACA 2580
OY 2581 GATGGGAGTTCGATTAAGATCAGATACGATACCAACACTACATTTTGACGACAAACCTT 2640
DB 2581 GATGGGAGTTCGATTAAGATCAGATACGATACCAACACTACATTTTGACGACAAACCTT 2640
OY 2641 CTTCGAGATACAGCAAGAAACAGCAGCAACAGTCTTCAGGTACTCCGCGCTGACAGGAG 2700
DB 2641 CTTCGAGATACAGCAAGAAACAGCAGCAACAGTCTTCAGGTACTCCGCGCTGACAGGAG 2700
OY 2701 CCTTGA 2706
DB 2701 CCTTGA 2706
```

```
RESULT 2
US-09-920-705-1
: Sequence 1, Application US/09920705
: Publication No. US20030079252A1
: GENERAL INFORMATION:
: APPLICANT: Amasino, Richard M.
: APPLICANT: Schomburg, Fritz M.
: APPLICANT: Michaels, Scott D.
: APPLICANT: Patton, David
: TITLE OF INVENTION: Floral Induction Gene
: FILE REFERENCE: 960296.97214
: CURRENT APPLICATION NUMBER: US/09/920,705
: CURRENT FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4593
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-920-705-1
```

```
Query Match 70.1%; Score 1896.8; DB 9; Length 4593;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2137; Conservative 0; Mismatches 2; Indels 229; Gaps 1;
```

```
OY 568 CAGTATCCTCCTCATATGAAGACTTTAAAGAGATGTCAGCAAGTAAGGTTCTGTGG 627
DB 2226 CAGTATCCTCCTCATATGAAGACTTTAAAGAGATGTCAGCAAGTAAGGTTCTGTGG 2285
OY 628 ATTGGGTTCCCTCTACTGCTACACATGCAATGATGAGCAAAATTCGCAAAATGCGATG 687
DB 2286 ATTGGGTTCCCTCTACTGCTACACATGCAATGATGAGCAAAATTCGCAAAATGCGATG 2345
OY 688 ATACTCTTGTGATGAGAGGTTAAAGATTACCATCAAGAAATTTTGCACTGTG 747
DB 2346 ATACTCTTGTGATGAGAGGTTAAAGATTACCATCAAGAAATTTTGCACTGTG 2405
OY 748 GAGTTTAGAGCGCGAGGAAGCTGCGCAATGCAAGGAAGGCTTACAGGGAGTTATTC 807
DB 748 GAGTTTAGAGCGCGAGGAAGCTGCGCAATGCAAGGAAGGCTTACAGGGAGTTATTC 807
```

```
DB 2406 GAGTTTAGAGCGCGAGGAAGCTGCGCAATGCAAGGAAGGCTTACAGGGAGTTATTC 2465
OY 808 AATAATCCTAGATCAAAATTAATGATGCTCAAAAGATGAGTTGCTCCTGAGCAAGAGCAT 867
DB 2466 AATAATCCTAGATCAAAATTAATGATGCTCAAAAGATGAGTTGCTCCTGAGCAAGAGCAT 2525
OY 868 ACTAGTTTCTGCTGATGAGAAAGGTCGACGAGATGATGTTCAATATGATGCTTCA 927
DB 2526 ACTAGTTTCTGCTGATGAGAAAGGTCGACGAGATGATGTTCAATATGATGCTTCA 2585
OY 928 TGTGATCTTCTCCTCATTTCTACTGGAATTCCTGGGTATGAGGCCCTCAGAGTACG 987
DB 2586 TGTGATCTTCTCCTCATTTCTACTGGAATTCCTGGGTATGAGGCCCTCAGAGTACG 2645
OY 988 AATGAGGCTTATATATGTCGACGATATCAATGATGCTGTTGGTAAAGGACCAAACTGG 1047
DB 2646 AATGAGGCTTATATATGTCGACGATATCAATGATGCTGTTGGTAAAGGACCAAACTGG 2705
OY 1048 AGGAGGCCATTCGCAAAATGGAACGTAATCTCCCATCTCCACAGACGCTGGAATCTC 1107
DB 2706 AGGAGGCCATTCGCAAAATGGAACGTAATCTCCCATCTCCACAGACGCTGGAATCTC 2765
OY 1108 CCATCTCCTGACCAAGATGACGAGGCCCTATGAGGTCAAAACCCGATTTCTGGGAAGA 1167
DB 2766 CCATCTCCTGACCAAGATGACGAGGCCCTATGAGGTCAAAACCCGATTTCTGGGAAGA 2825
OY 1168 TATGATCTGCTGATGTCGATGAGAAAGTAAAGAACCAAGAGATGATGATGATGATG 1227
DB 2826 TATGATCTGCTGATGTCGATGAGAAAGTAAAGAACCAAGAGATGATGATGATGATG 2885
OY 1228 GGTTTTACTCCAAATGGGTCGATGAGAGGTCATTTGGTCGAGGTCAGTTGCTGCTAGA 1287
DB 2886 GGTTTTACTCCAAATGGGTCGATGAGAGGTCATTTGGTCGAGGTCAGTTGCTGCTAGA 2945
OY 1288 CCTATCGGTGCCCCCTGATTCGATCAATGATGAGAGGATGATGATGATGATGATGATG 1347
DB 2946 CCTATCGGTGCCCCCTGATTCGATCAATGATGAGAGGATGATGATGATGATGATGATG 3005
OY 1348 ACTCCGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1403
DB 3006 ACTCCGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3065
OY 1404 ----- 1403
DB 3066 GFACTAATTTCTAGACCTTTAACCTTCTAGTGTTCCTTTTCAGACGATTTATATAT 3125
OY 1404 ----- 1403
DB 3126 TTTCATTTCAATTCGATGAGAGTAACATTTATAGATGATACATTTTATTTACTA 3185
OY 1404 ----- 1403
DB 3186 TTACTGTGTAGTTCGAGATGTCGTCGATTTTCATGATGATGATGATGATGATGATG 3245
OY 1404 ----- 1403
DB 3246 CCCCAATTTACTGACTTTGTTTTTTTTTTTAAATGATTTATAGGCTGAGGTCGTAA 3305
OY 1419 TTGTTACAGCAAGATGATTTGAATGATGCTGCTAAACATTTAGCCGTTGCAATGATG 1478
DB 3306 TTGTTACAGCAAGATGATTTGAATGATGCTGCTAAACATTTAGCCGTTGCAATGATG 3365
OY 1479 TGAGATGCTTTTTCGATGACAGACAGGAGGATTTGCGCTTACATGATGATGATGATG 1538
DB 3366 TGAGATGCTTTTTCGATGACAGACAGGAGGATTTGCGCTTACATGATGATGATGATG 3425
OY 1539 CCGGTACTTGTAGCTCAAAAGATCGGCGGCTGTCGCAAAATGATGATGATGATGATG 1598
DB 3426 CCGGTACTTGTAGCTCAAAAGATCGGCGGCTGTCGCAAAATGATGATGATGATGATG 3485
OY 1599 ATTCTTGTGCTCCTCATGATGATTTCTTATGATGATGATGATGATGATGATGATG 1658
DB 3486 ATTCTTGTGCTCCTCATGATGATTTCTTATGATGATGATGATGATGATGATGATG 3545
```

```
OY 1659 GCTATATGTTGTTCTCAAGTTACCCCGCAGCCGTTCTGTATACAGATCATACAG 1718
    |||||
Db 3546 GCTATATGTTGTTCTCAAGTTACCCCGCAGCCGTTCTGTATACAGATCATACAG 3605
OY 1719 ACAGATATCTAGTCCAAATCTCTGCATTAATATGATCAGCCCGGATTCACCTGCCAA 1778
    |||||
Db 3606 ACAGATATCTAGTCCAAATCTCTGCATTAATATGATCAGCCCGGATTCACCTGCCAA 3665
OY 1779 TGCTAGTACAGTTATCTCTCTAGGAAATTAATCTATAGGGGTGCACCAAGACATTT 1838
    |||||
Db 3666 TGCTAGTACAGTTATCTCTCTAGGAAATTAATCTATAGGGGTGCACCAAGACATTT 3725
OY 1839 GACAGCTGCTTCAAAACCATCTGTTAGGAGACCTCTGAGATACCTATATAATGACGGCC 1898
    |||||
Db 3726 GACAGCTGCTTCAAAACCATCTGTTAGGAGACCTCTGAGATACCTATATAATGACGGCC 3785
OY 1899 TCAAGCTGGGGTGTATTAATCTCGAGACTTTTAGCCACTCTGACATCTATCTCCCTGC 1958
    |||||
Db 3786 TCAAGCTGGGGTGTATTAATCTCGAGACTTTTAGCCACTCTGACATCTATCTCCCTGC 3845
OY 1959 AACTTCTCAACCTGCTGCCCCGAGAGACACCACTATGTCAGAGACCTTCAACAGTTGT 2018
    |||||
Db 3846 AACTTCTCAACCTGCTGCCCCGAGAGACACCACTATGTCAGAGACCTTCAACAGTTGT 3905
OY 2019 TTCCACAGACATCATGTCATGAGTGTATCAATGAGAAACACCTCTCAAGCTTGAA 2078
    |||||
Db 3906 TTCCACAGACATCATGTCATGAGTGTATCAATGAGAAACACCTCTCAAGCTTGAA 3965
OY 2079 AAGAGGTCCCAAAACAGTTATGATGCTCAATCACTCATCTCCACAAATGCGAAATCA 2138
    |||||
Db 3966 AAGAGGTCCCAAAACAGTTATGATGCTCAATCACTCATCTCCACAAATGCGAAATCA 4025
OY 2139 GTACACTCCAGCTGGGCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2198
    |||||
Db 4026 GTACACTCCAGCTGGGCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4085
OY 2199 CCCCACATCACTATGTTGAATGTCATGTCACATGCAATGACCAAGCCAACTCTGTAA 2258
    |||||
Db 4086 CCCCACATCACTATGTTGAATGTCATGTCACATGCAATGACCAAGCCAACTCTGTAA 4145
OY 2259 CATGCTTCAGCTGCTCCGTTTACCAAAATATGCTCAATTAATTAATTCATCTGACCTCA 2318
    |||||
Db 4146 CATGCTTCAGCTGCTCCGTTTACCAAAATATGCTCAATTAATTAATTCATCTGACCTCA 4205
OY 2319 GGGTTCGCAAAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2378
    |||||
Db 4206 GGGTTCGCAAAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4265
OY 2379 CATGCTTCAGCTGCTCCGTTTACCAAAATATGCTCAATTAATTAATTCATCTGACCTCA 2438
    |||||
Db 4266 CATGCTTCAGCTGCTCCGTTTACCAAAATATGCTCAATTAATTAATTCATCTGACCTCA 4325
OY 2439 AACACAAATCAGGACAGAGAACTTAAACCTTCCCAATTCAGCTGCGCATGCAACACAC 2498
    |||||
Db 4326 AACACAAATCAGGACAGAGAACTTAAACCTTCCCAATTCAGCTGCGCATGCAACACAC 4385
OY 2499 AGCAGATAGGCAAAATTTAGAGCCACAAACCAAGCACTACGATTCAGCTTATGATCTC 2558
    |||||
Db 4386 AGCAGATAGGCAAAATTTAGAGCCACAAACCAAGCACTACGATTCAGCTTATGATCTC 4445
OY 2559 TGGGGATGCTCAGGGGTACAAACAGATGGGAGTCTGATTAAGATTCAGAGATACAGTCAAC 2618
    |||||
Db 4446 TGGGGATGCTCAGGGGTACAAACAGATGGGAGTCTGATTAAGATTCAGAGATACAGTCAAC 4505
OY 2619 ACTACAAATTTGACAGAAACCTTCTCTCAGATATACAGAGAAACAGACAGACAGTCTTC 2678
    |||||
Db 4506 ACTACAAATTTGACAGAAACCTTCTCTCAGATATACAGAGAAACAGACAGACAGTCTTC 4565
OY 2679 AGGTACTCCGGCTGACAGAGGGGCTTGA 2706
    |||||
Db 4566 AGGTACTCCGGCTGACAGAGGGGCTTGA 4593
```

```
RESULT 3
US-09-920-705-4/c
; Sequence 4, Application US/09920705
; Publication No. US20030079252A1
; GENERAL INFORMATION:
; APPLICANT: Amasino, Richard M.
; APPLICANT: Schomburg, Scott D.
; APPLICANT: Michaels, Fritz D.
; APPLICANT: Patton, David
; TITLE OF INVENTION: Floral Induction Gene
; FILE REFERENCE: 960296.97214
; CURRENT APPLICATION NUMBER: US/09/920,705
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 4
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; OTHER INFORMATION: FPA antisense fragment
US-09-920-705-4
```

```
Query Match          29.5%; Score 797.8; DB 9; Length 801;
Best Local Similarity 99.8%; Pred. No. 2.1e-249;
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 587 AAGACTTTAAAGAGATGTCACGCAAGATTAAGTTGTGATTTGGTTCCCTCTACTG 646
    |||||
Db 801 AAGACTTTAAAGAGATGTCACGCAAGATTAAGTTGTGATTTGGTTCCCTCTACTG 742
OY 647 CTACACAAATCATGATGAGCAAAATCTGCACAAATCGATGATCTCTTTGGTAGATG 706
    |||||
Db 741 CTACACAAATCATGATGAGCAAAATCTGCACAAATCGATGATCTCTTTGGTAGATG 682
OY 707 AGAGGTTAAAGATGTCACGCAAGATTTGACATTTGAGTTAGAGCGGAGG 766
    |||||
Db 681 AGAGGTTAAAGATTTACCATCAAGAAATTTTGCATTTGTGAGTTAGAGCGGAGG 622
OY 767 AAGCTGCCAATTCAGAGAAAGGCTTACAGGGAGGTTATTCATAATCTCTAGAAATCAAA 826
    |||||
Db 621 AAGCTGCCAATTCAGAGAAAGGCTTACAGGGAGGTTATTCATAATCTCTAGAAATCAAA 562
OY 827 TTAATGATCAAAAGATGTCGCTGCTGAGCAAGATTAATGTTTACTCTGTA 886
    |||||
Db 561 TTAATGATCAAAAGATGTCGCTGCTGAGCAAGATTAATGTTTACTCTGTA 502
OY 887 TGAAGGTCAGAGCAGATATGTTCAATATGATCTCATGTGATCTCTCCATTT 946
    |||||
Db 501 TGAAGGTCAGAGCAGATATGTTCAATATGATCTCATGTGATCTCTCCATTT 442
OY 947 CTACGGAATTTCTGGGTCTATGAGGCCCCCTAGAGATTCGAATGAGCTTATTAATG 1006
    |||||
Db 441 CTACGGAATTTCTGGGTCTATGAGGCCCCCTAGAGATTCGAATGAGCTTATTAATG 382
OY 1007 GTGCAATATACATGACGTTGTTGTAAGAGCCAAACGTGAGAGGCGCATGTGAAATG 1066
    |||||
Db 381 GTGCAATATACATGACGTTGTTGTAAGAGCCAAACGTGAGAGGCGCATGTGAAATG 322
OY 1067 GAAGTGAATTAATCTCCATCTCCACAGAGACCTGGAATCTCCATCTCTGACAAAGTA 1126
    |||||
Db 321 GAAGTGAATTAATCTCCATCTCCACAGAGACCTGGAATCTCCATCTCTGACAAAGTA 262
OY 1127 CGAGGGGCCCTATGAGGTCAAACCCGANTCTTGGGAAGATATATCTCTGCTCAGTTGG 1186
    |||||
Db 261 CGAGGGGCCCTATGAGGTCAAACCCGANTCTTGGGAAGATATATCTCTGCTCAGTTGG 202
OY 1187 TCAGAGAAAGTAAAGCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
    |||||
Db 201 TCAGAGAAAGTAAAGCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 142
OY 1247 TCAGATGAGAGTCAATTTTGTGAGGTTCAAGTTGCTGTAAGACATATCCGTTGCCCCCTGT 1306
```



Db 141 TCGATGAGAGCGCATTTGGTTCAGAGTTCAGTTCGCTGACCTATCCCTGGCCCCCTG 82  
Qy 1307 ATTGATACATATGAGAGGAATGATCCCAAGGGTGAACCTCCGCTCTGTGCTC 1366  
Db 81 ATTCTGATCACATATGAGAGGAATGATCCCAAGGGTGAACCTCCGCTCTGTGCTC 22  
Qy 1367 GTTGTGTACTATGGGAAG 1387  
Db 21 GTTGTGTACTATGGGAAG 1

RESULT 4  
US-09-920-705-5

Sequence 5, Application US/09920705  
Publication No. US20030079252A1  
GENERAL INFORMATION:  
APPLICANT: Amasino, Richard M.  
APPLICANT: Schomburg, Fritz M.  
APPLICANT: Michaels, Scott D.  
APPLICANT: Paton, David  
TITLE OF INVENTION: Floral Induction Gene  
FILE REFERENCE: 960296, 97214  
CURRENT APPLICATION NUMBER: US/09/920,705  
CURRENT FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 349  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: Portion of FPA coding region  
US-09-920-705-5

Query Match 10.0%; Score 271.4; DB 9; Length 349;  
Best Local Similarity 94.6%; Pred. No. 1.1e-77;  
Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 ATGGCGTTATCTATGAGGCAATTCAGAGCGGATGATCCGGTTCCAGTCAACAATCTT 60  
Db 1 ATGGCGTTATCTATGAGGCAATTCAGAGCGGATGATCCGGTTCCAGTCAACAATCTT 60  
Qy 61 TGGGTCGTAGCGCTAACGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
Db 61 TGGGTCGTAGCGCTAACGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
Qy 121 TACGGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
Db 121 TACGGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
Qy 181 AGACATGTGAG 240  
Db 181 AGACATGTGAG 240  
Qy 241 AGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297  
Db 241 AGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297

RESULT 5  
US-09-920-705-6

Sequence 6, Application US/09920705  
Publication No. US20030079252A1  
GENERAL INFORMATION:  
APPLICANT: Amasino, Richard M.  
APPLICANT: Schomburg, Fritz M.  
APPLICANT: Michaels, Scott D.  
APPLICANT: Paton, David  
TITLE OF INVENTION: Floral Induction Gene  
FILE REFERENCE: 960296, 97214  
CURRENT APPLICATION NUMBER: US/09/920,705  
CURRENT FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 3715  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: FPA promoter plus Intron  
US-09-920-705-6

Query Match 10.0%; Score 271.4; DB 9; Length 3715;  
Best Local Similarity 94.6%; Pred. No. 6.8e-77;  
Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 ATGGCGTTATCTATGAGGCAATTCAGAGCGGATGATCCGGTTCCAGTCAACAATCTT 60  
Db 1832 ATGGCGTTATCTATGAGGCAATTCAGAGCGGATGATCCGGTTCCAGTCAACAATCTT 1891  
Qy 61 TGGGTCGTAGCGCTAACGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
Db 1892 TGGGTCGTAGCGCTAACGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1951  
Qy 121 TACGGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180  
Db 1952 TACGGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2011  
Qy 181 AGACATGTGAG 240  
Db 2012 AGACATGTGAG 2071  
Qy 241 AGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297  
Db 2072 AGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2128

RESULT 6  
US-09-770-444-972

Sequence 972, Application US/09770444  
Patent No. US2002023280A1  
GENERAL INFORMATION:  
APPLICANT: Gorlach, Jörn  
APPLICANT: An, Yong-Olang  
APPLICANT: Hamilton, Carol M.  
APPLICANT: Price, Jennifer L.  
APPLICANT: Raines, Tracy M.  
APPLICANT: Yu, Yang  
APPLICANT: Rameeka, Joshua G.  
APPLICANT: Page, Amy  
APPLICANT: Matthew, Abraham V.  
APPLICANT: Ledford, Brooke L.  
APPLICANT: Woessner, Jeffrey P.  
APPLICANT: Haas, William David  
APPLICANT: Garcia, Carlos A.  
APPLICANT: Krickler, Maja  
APPLICANT: Slader, Ted  
APPLICANT: Davis, Keith R.  
APPLICANT: Allen, Keith  
APPLICANT: Hoffman, Neil  
APPLICANT: Hurban, Patrick  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2027 (PARA-016PRV)  
CURRENT APPLICATION NUMBER: US/09/770,444  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,502  
NUMBER OF SEQ ID NOS: 999  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 972  
LENGTH: 438  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-770-444-972

```

      APPLICANT : Elashoff, Michael
      APPLICANT : Gene Logic, Inc.
      TITLE OF INVENTION: Molecular Toxicology Modeling
      FILE REFERENCE: 44921-5038-US
      CURRENT FILING DATE: US/09/917,800A
      PRIOR APPLICATION NUMBER: 2001-07-31
      PRIOR FILING DATE: 2000-07-31
      PRIOR APPLICATION NUMBER: US 60/222,040
      PRIOR FILING DATE: 2000-11-02
      PRIOR APPLICATION NUMBER: US 60/290,029
      PRIOR FILING DATE: 2001-05-11
      PRIOR APPLICATION NUMBER: US 60/290,645
      PRIOR FILING DATE: 2001-05-15
      PRIOR APPLICATION NUMBER: US 60/292,336
      PRIOR FILING DATE: 2001-05-22
      PRIOR APPLICATION NUMBER: US 60/295,798
      PRIOR FILING DATE: 2001-06-06
      PRIOR APPLICATION NUMBER: US 60/297,457
      PRIOR FILING DATE: 2001-06-13
      PRIOR APPLICATION NUMBER: US 60/298,884
      PRIOR FILING DATE: 2001-06-19
      PRIOR APPLICATION NUMBER: US 60/303,459
      PRIOR FILING DATE: 2001-07-09
      NUMBER OF SEQ ID NOS: 1740
      SOFTWARE: Patentin Ver. 2.1
      SEQ ID NO 1567
      LENGTH: 2142
      TYPE: DNA
      ORGANISM: Rattus norvegicus
      FEATURE:
      OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012749
      US-09-917-800A-1567

Query Match          ,1.5%; Score 41.6; DB 10; Length 2142;
Best Local Similarity 59.2%; Pident.No. 0.055;
Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY    150 TTCTTCACGAGCGCTTTCGGTTATATACTCAGACATGTGGAGGAGCGAGCCAA 209
      |||||   || |||||   || | || | || || || || || || || || || ||
Db    1830 TTCTTCMAAGGGGTTTGTTGTGATGACTTAAATAGGAGGAAGATGCCAAAGCTGCCAA 1889
      |||||   || | || | || || || || || || || || || || || || || ||

QY    210 AGAGCGCTCTCAAGAGCAACAATTGGATGGAGAAGTCAATTTAAGATCGCATAGCACGACC 269
      |||||   || | || | || || || || || || || || || || || || || ||
Db    1890 GGAGCGCATGAGAGATGAGAAATTGATGAGAACAAAGTTACTTGGACTGGCCCAATCC 1949
      |||||   || | || | || || || || || || || || || || || || || ||

RESULT 9
US-10-123-155-438
Sequence 438, Application US/10123155
Publication No. US20030068794A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155

```

APPLICANT:	Cahoon, Edgar B
APPLICANT:	Cahoon, Rebecca E
APPLICANT:	Palco, Savvelto Carr
APPLICANT:	Feng, Ylven
APPLICANT:	Hantke, Sabine S.
APPLICANT:	Lee, Jian-Ming
APPLICANT:	Li, Zhongsen
APPLICANT:	Mao, Guo-Hua
APPLICANT:	Morgan, Michele
APPLICANT:	Miu, Xiling
APPLICANT:	Ogell, Joan
APPLICANT:	Rafalski, Antoni
APPLICANT:	Sekali, Hajime
APPLICANT:	Zheng, Feizhong

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresinski, Maureen  
 APPLICANT: Deforge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel

Query Match	1.4%;	Score 38.2;	DB 9;	Length 1049;
Best Local Similarity	5.7%;	Pred. No. 0.41;		
Matches	40;	Conservative 176;	Mismatches 483;	Indels 0;
				Gaps 0;

RESULT 12  
US-09-878-574-4075  
; Sequence 4075, Application US/09878574  
; Patent No. US20020110548A1

Query Match	1.4%;	Score 37.2;	DB 10;	Length 424;
Best Local Similarity	61.2%;	Pred. No. 0.45;		
Matches 60;	Conservative 0;	Mismatches 38;	Indels 0;	Gaps 0

RESULT 13  
US-10-062-254-275  
; Sequence 275, Application US/10062254

GENERAL INFORMATION:  
 APPLICANT: Cahoon, Edgar B  
 APPLICANT: Cahoon, Rebecca E  
 APPLICANT: Falco, Saverio Carl  
 APPLICANT: Fang, Yiwen  
 APPLICANT: Hantke, Sabine S.  
 APPLICANT: Lee, Jian-Ming  
 APPLICANT: Li, Zhongsen  
 APPLICANT: Miao, Guo-Hua  
 APPLICANT: Morgante, Michele  
 APPLICANT: Niu, Xiping  
 APPLICANT: Ogelli, Joan  
 APPLICANT: Rafalski, Antoni  
 APPLICANT: Sakai, Hajime  
 APPLICANT: Zheng, Peizhong  
 APPLICANT: Zhu, Qun  
 FILE REFERENCE:  
 TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism  
 CURRENT APPLICATION NUMBER: US/10/062,254  
 CURRENT FILING DATE: 2002-02-01  
 PRIOR APPLICATION NUMBER: 09/630,346  
 PRIOR FILING DATE: 2000-07-28  
 PRIOR APPLICATION NUMBER: 60/146511  
 PRIOR FILING DATE: 1999-07-30  
 PRIOR APPLICATION NUMBER: 60/156006  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: 60/156899  
 PRIOR FILING DATE: 1999-09-30  
 PRIOR APPLICATION NUMBER: 60/157287  
 PRIOR FILING DATE: 1999-10-01  
 PRIOR APPLICATION NUMBER: 60/169767  
 PRIOR FILING DATE: 1999-12-09  
 PRIOR APPLICATION NUMBER: 60/171054  
 PRIOR FILING DATE: 1999-12-16  
 PRIOR APPLICATION NUMBER: 60/172958  
 PRIOR FILING DATE: 1999-12-21  
 PRIOR APPLICATION NUMBER: 60/171515

;; PRIOR FILING DATE: 1999-12-22  
;; PRIOR APPLICATION NUMBER: 60/173535  
;; PRIOR FILING DATE: 1999-12-29  
;; NUMBER OF SEQ ID NOS: 375  
;; SOFTWARE: Microsoft Office 97  
;; SEQ ID NO 275  
;; LENGTH: 1537  
;; TYPE: DNA  
;; ORGANISM: Glycine max  
US-10-062-254-275

Query Match 1.4%; Score 37.2; DB 12; Length 1537;  
Best Local Similarity 61.2%; Pred. No. 1.2;  
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

OY 282 TAAGAGTCTATGAGGCGGCGGCTTAATGCTCCAGAGATGAGCGAGAGA 341  
DB 879 TACACTATTTTGTGTGTAATTTGATCTTAATGCTCATGATCATCTGAGCGAAGT 938  
OY 342 GTTCAGCAAGTTTGGGAAATCGAGATTATTAGGTTTC 379  
DB 939 TTTCAGCCAGTATGAGATTAATGATGATGGAAGATTC 976

## RESULT 14

US-10-197-666A-77  
;; Sequence 77, Application US/10197666A  
;; Publication No. US20030092037A1  
;; GENERAL INFORMATION:  
;; APPLICANT: ASAH KASEI KABUSIKI KAISYA  
;; TITLE OF INVENTION: ELKI phosphorylation related gene  
;; FILE REFERENCE: PH-1548US  
;; CURRENT APPLICATION NUMBER: US/10/197,666A  
;; PRIOR FILING DATE: 2002-11-18  
;; PRIOR APPLICATION NUMBER: JP 2001-218204  
;; PRIOR FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: JP 2001-263450  
;; PRIOR FILING DATE: 2001-08-31  
;; PRIOR APPLICATION NUMBER: JP 2002-012176  
;; PRIOR FILING DATE: 2002-01-21  
;; PRIOR APPLICATION NUMBER: US 60/305,884  
;; PRIOR FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: US 60/316,304  
;; PRIOR FILING DATE: 2001-09-04  
;; PRIOR APPLICATION NUMBER: US 60/350,027  
;; PRIOR FILING DATE: 2002-01-23  
;; NUMBER OF SEQ ID NOS: 156  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 77  
;; LENGTH: 1846  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (128)..(877)  
US-10-197-666A-77

Query Match 1.4%; Score 37; DB 9; Length 1846;  
Best Local Similarity 54.9%; Pred. No. 1.6;  
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 153 TTCAGAGGCTTTGCTTATATCTACAGACATGAGAGAGCGACGCCAAGA 212  
DB 484 TTCAAGAGAGATTGCTTGTATATTGAAAATGTAGATGCGCAAGAGCTAAGA 543  
OY 213 GGCCTTCAAGAGCAAAATTTGAATGAATGCAATTAAGATCGACGACCGGC 272  
DB 544 ACGTGCCAATGGAATGAGCTTGATGGCGGTAGATCAGAGTTTCTATATACAAA 603  
OY 273 AAAACCTTGTAAAG 285  
DB 604 AAGACCACTATACG 616

## RESULT 15

US-10-197-666A-79  
;; Sequence 79, Application US/10197666A  
;; Publication No. US20030092037A1  
;; GENERAL INFORMATION:  
;; APPLICANT: ASAH KASEI KABUSIKI KAISYA  
;; TITLE OF INVENTION: ELKI phosphorylation related gene  
;; FILE REFERENCE: PH-1548US  
;; CURRENT APPLICATION NUMBER: US/10/197,666A  
;; PRIOR FILING DATE: 2002-11-18  
;; PRIOR APPLICATION NUMBER: JP 2001-218204  
;; PRIOR FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: JP 2001-263450  
;; PRIOR FILING DATE: 2001-08-31  
;; PRIOR APPLICATION NUMBER: JP 2002-012176  
;; PRIOR FILING DATE: 2002-01-21  
;; PRIOR APPLICATION NUMBER: US 60/305,884  
;; PRIOR FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: US 60/316,304  
;; PRIOR FILING DATE: 2001-09-04  
;; PRIOR APPLICATION NUMBER: US 60/350,027  
;; PRIOR FILING DATE: 2002-01-23  
;; NUMBER OF SEQ ID NOS: 156  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 79  
;; LENGTH: 1964  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (130)..(993)  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: 38  
;; OTHER INFORMATION: n=A or C or G or T  
US-10-197-666A-79

Query Match 1.4%; Score 37; DB 9; Length 1964;  
Best Local Similarity 54.9%; Pred. No. 1.6;  
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 153 TTCAGAGGCTTTGCTTATATCTACAGACATGAGAGAGCGACGCCAAGA 212  
DB 600 TTCAAGAGAGATTGCTTGTATATTGAAAATGTAGATGCGCAAGAGCTAAGA 659  
OY 213 GGCCTTCAAGAGCAAAATTTGAATGAATGCAATTAAGATCGAATCGACGACCGGC 272  
DB 660 ACGTGCCAATGGAATGAGCTTGATGGCGGTAGATCAGAGTTTCTATATACAAA 719  
OY 273 AAAACCTTGTAAAG 285  
DB 720 AAGACCACTATACG 732

Search completed: June 19, 2003, 07:32:24  
Job time : 933 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 01:11:43 ; Search time 128 Seconds

(without alignments)  
6483.338 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706  
Sequence: 1 atggcgtatcatatgaagcc.....cggctgacagggccttga 2706

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query Length	ID	Description
C 1	47.6	1.8	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	35.4	1.3	3489	2 US-08-728-323A-1	Sequence 1, Appl
C 3	35.4	1.3	3489	4 US-09-298-568-1	Sequence 1, Appl
C 4	35.4	1.3	32207	2 US-08-770-379-20	Sequence 20, Appl
C 5	35.4	1.3	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 6	35.4	1.3	32207	4 US-09-230-371A-20	Sequence 5, Appl
C 7	34.8	1.3	2406	4 US-09-632-098-5	Sequence 6, Appl
C 8	34.8	1.3	2439	4 US-09-632-098-6	Sequence 3, Appl
C 9	34.6	1.3	561	2 US-08-589-711-3	Sequence 22, Appl
C 10	34.6	1.3	561	4 US-09-117-257-22	Sequence 22, Appl
C 11	34.6	1.3	561	4 US-09-117-257-24	Sequence 24, Appl
C 12	34.6	1.3	561	4 US-09-221-938-3	Sequence 8, Appl
C 13	34.6	1.3	561	4 US-08-945-476-8	Sequence 22, Appl
C 14	34.6	1.3	561	4 US-08-945-476-22	Sequence 22, Appl
C 15	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 16	34.6	1.3	561	4 US-09-489-352-22	Sequence 22, Appl
C 17	34.6	1.3	561	4 US-09-489-352-24	Sequence 24, Appl
C 18	34.6	1.3	561	4 US-09-117-257-29	Sequence 29, Appl
C 19	34.6	1.3	561	4 US-09-117-257-43	Sequence 29, Appl
C 20	34.6	1.3	564	4 US-09-489-352-29	Sequence 29, Appl
C 21	34.6	1.3	564	4 US-09-489-352-43	Sequence 43, Appl
C 22	34.6	1.3	1253	4 US-09-117-257-11	Sequence 11, Appl
C 23	34.6	1.3	1253	4 US-08-945-476-11	Sequence 11, Appl
C 24	34.6	1.3	1253	4 US-09-489-352-11	Sequence 11, Appl
C 25	34.6	1.3	2653	2 US-08-589-711-1	Sequence 1, Appl
C 26	34.6	1.3	2653	4 US-09-221-938-1	Sequence 7, Appl
C 27	34.6	1.3	2653	4 US-08-945-476-7	Sequence 7, Appl

C 28	34.6	1.3	2656	4 US-09-117-257-7	Sequence 7, Appl
C 29	34.6	1.3	2656	4 US-09-489-352-7	Sequence 9, Appl
C 30	34.2	1.3	2849	4 US-09-221-017B-990	Sequence 990, App
C 31	34	1.3	1851	3 US-09-042-426-3	Sequence 3, Appl
C 32	34	1.3	1851	4 US-09-291-238-3	Sequence 3, Appl
C 33	34	1.3	1851	4 US-09-330-760-3	Sequence 3, Appl
C 34	34	1.3	1851	4 US-09-328-473-3	Sequence 3, Appl
C 35	34	1.3	1851	4 US-09-330-737-3	Sequence 3, Appl
C 36	34	1.3	1851	4 US-09-329-169-3	Sequence 3, Appl
C 37	34	1.3	1851	4 US-09-330-714A-3	Sequence 3, Appl
C 38	34	1.3	1851	4 US-09-328-826-3	Sequence 3, Appl
C 39	34	1.3	7378	3 US-09-042-426-9	Sequence 9, Appl
C 40	34	1.3	7378	4 US-09-291-228-9	Sequence 9, Appl
C 41	34	1.3	7378	4 US-09-330-760-9	Sequence 9, Appl
C 42	34	1.3	7378	4 US-09-328-473-9	Sequence 9, Appl
C 43	34	1.3	7378	4 US-09-330-737-9	Sequence 9, Appl
C 44	34	1.3	7378	4 US-09-329-169-9	Sequence 9, Appl
C 45	34	1.3	7378	4 US-09-330-714A-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEFFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOMIPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELE: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
; US-08-232-463-14  
Query Match 1.8%, Score 47.6, DB 1, Length 7218;



Best Local Similarity 6.8%; Pred. No. 0.00015;  
Matches 29; Conservative 214; Mismatches 183; Indels 0; Gaps 0;

QY 132 TGATAGATCAGCGTGTCTTCTCAGCGAGCGTTCGCTTTATTAAGTACAGATGCGA 191  
 Db 1496 TGTAGGCTCTACTGTAATCTATCTATGCAAGTATTAAGAGATGAGATTTGGT 1437  
 QY 192 GGAAGCAGTCCAGCCAAAGAGCGCTTTCAGAGCAAAATTTGAATGGAAGTCAATTA 251  
 Db 1436 ACRRR 1377  
 QY 252 GATGCAATGCGCAGCAGCGCAAAACCTTTGAGAGCTTATGGGTGGGATCGGCC 311  
 Db 1376 RRR 1317  
 QY 312 TAATGCTCCAGAGTGCCTGGAGAGAGTTCAGAGATTTGGGAAATTCAGAGATT 371  
 Db 1316 RRR 1257  
 QY 372 TAGGTTCTCAGAGACGAGAGAGCTTTCATTGATTAATGAGATGATGATGCTT 431  
 Db 1256 RRR 1197  
 QY 432 ACAGCTAAGAGCATGATGAAGCCTATGGGTGAGCTTTTGCTGTGATTCT 491  
 Db 1196 RRR 1137  
 QY 492 CCGGTCAACAGCGCCAAAAGAAACAATGGGCTGCTTACGATTAAGAAATGGCAA 551  
 Db 1136 RRR 1077  
 QY 552 TATGAA 557  
 Db 1076 RRRRRR 1071

## RESULT 2

US-08-728-323A-1/c  
 ; Sequence 1, Application US/08728323A  
 ; Patent No. 5948676

## GENERAL INFORMATION:

APPLICANT: Chang, Yuan  
 APPLICANT: Bohenzky, Roy A.  
 APPLICANT: Russo, James J.  
 APPLICANT: Edelman, Isidore S.  
 APPLICANT: Moore, Patrick S.  
 TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
 TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
 TITLE OF INVENTION: Encoding Same And Uses Thereof  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/728,323A  
 FILING DATE:

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSK/SKS  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400  
 TELEFAX: 212-391-0525

## INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 3489 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS  
 LOCATION: 1..3489

US-08-728-323A-1

## Query Match

Best Local Similarity 1.3%; Score 35.4; DB 2; Length 3489;  
 Matches 147; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 1650 TCAAGAGCGCTATATGTTGTTCTTCAAGTACCCCGCAGCGCTTCTTACAGC 1709  
 Db 2381 TCCTGCTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2232  
 QY 1710 ATCATACAGACAGAAATCTCAGTCCATCTCTGCAATATATGATCAAGCCGGATTC 1769  
 Db 2321 TGTCTCTGCTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2262  
 QY 1770 ACCGCAATAGTACGACAGTTATATCTCTCTAGGAAATATACATTAGGGGTGACCC 1829  
 Db 2261 TCCTGCTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2202  
 QY 1830 AGACATTTACAGCTGCTTCAAAACATCTGTTAGCGAGCTCTAGAAATCTAATA 1889  
 Db 2201 TCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2142  
 QY 1890 TGCAGCGCTCAGCTGGGTTAGTTTAACCTCGGAGCTTTAGCCAGCTGGCACTAT 1949  
 Db 2141 TGTCTGCTCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2082  
 QY 1950 TCTCCTGCACTTCTCAACCTGCTGCCCTGA 1982  
 Db 2081 TCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2049

## RESULT 3

US-09-298-568-1/c  
 ; Sequence 1, Application US/09298568  
 ; Patent No. 6322792

## GENERAL INFORMATION:

APPLICANT: Kieff, Elliott D.  
 APPLICANT: Balleskas, Mary E.  
 TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
 TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
 FILE REFERENCE: 16412-10001R  
 CURRENT APPLICATION NUMBER: US/09/298,568  
 CURRENT FILING DATE: 1999-04-21  
 EARLIER APPLICATION NUMBER: US 60/109,422  
 EARLIER FILING DATE: 1998-11-19  
 NUMBER OF SEQ ID NOS: 3  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 1  
 LENGTH: 3489

ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-1

## Query Match

Best Local Similarity 1.3%; Score 35.4; DB 4; Length 3489;  
 Matches 147; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 1650 TCAAGAGCGCTATATGTTGTTCTTCAAGTACCCCGCAGCGCTTCTTACAGC 1709  
 Db 2381 TCCTGCTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2232  
 QY 1710 ATCATACAGACAGAAATCTCAGTCCATCTCTGCAATATATGATCAAGCCGGATTC 1769

[illegible]

## RESULT 4

US-08-770-379-20  
; Sequence 20, Application US/08770379  
; Patent No. 5849564

## ; GENERAL INFORMATION:

APPLICANT: Chang, Yuan  
 APPLICANT: Bohenzky, Roy A.  
 APPLICANT: Russo, James J.  
 APPLICANT: Edelman, Isidore S.  
 APPLICANT: Moore, Patrick S.  
 TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
 TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
 NUMBER OF SEQUENCES: 20  
 \* CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/770,379  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 53342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 278-0400  
 TELEFAX: (212) 391-0525  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 32207 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-770-379-20

Query Match 1.3%; Score 35.4; DB 2; Length 32207;  
 Best Local Similarity 44.1%; Pred. No. 4;  
 Matches 147; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 1650 TCAGAACGGCTATATGCTGTGTTCTCAAGTACCCCGCCAGCCGCTTCCTTTCACAG 1709  
 |||||  
 Db 19616 TCTGCTCTCTATCTCTCTCTCTGCTCCCTTCAATCTCGCTCTGATCTCTTAACTATCG 19675  
 |||||

[illegible]

## RESULT 5

US-08-757-669A-20  
; Sequence 20, Application US/08757669A  
; Patent No. 6183751

## ; Patent No. 6183751

```

1 GENERAL INFORMATION:
2 APPLICANT: Chang, Yuan
3 APPLICANT: Bohenzky, Roy A.
4 APPLICANT: Russo, James J.
5 APPLICANT: Edelman, Isidore S.
6 APPLICANT: Moore, Patrick S.
7 TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
8 TITLE OF INVENTION: SEQUENCES AND USES THEREOF
9 NUMBER OF SEQUENCES: 20
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Cooper & Dunham LLP
12 STREET: 1185 Avenue of the Americas
13 CITY: New York
14 STATE: New York
15 COUNTRY: U.S.A.
16 ZIP: 10036
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.30
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/757,669A
25 FILING DATE:
26 CLASSIFICATION: 424
27 ATTORNEY/AGENT INFORMATION:
28 NAME: White, John P.
29 REGISTRATION NUMBER: 28,678
30 REFERENCE/DOCKET NUMBER: 45185-F
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (212) 278-0400
33 TELEFAX: (212) 391-0525
34 INFORMATION FOR SEQ ID NO: 20:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 32207 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: double
39 TOPOLOGY: linear
40 MOLECULE TYPE: DNA (genomic)
41
42 JS-08-757-669A-20

```

	Query Match	1.33;	Score 35.4;	DB 4;	Length 32207;
	Best Local Similarity	44.18;	Pred. No. 4;		
	Matches 14;	Conservative	0;	Mismatches 186;	Indels 0;
				Gaps 0;	
OY	1650	TCAGAACGGCTATATGTTGTTCTCAAGTTACCCCGCAGCGGTTCCGTACAC	1709		
Ob	19616	TCCTGCTCCCTACACCTCCTCTCTCTACCTCTCGTCTCGGCTCGGACCTCTACCTC	19675		

[illegible][illegible]

```

US-09-632-098-6
; Sequence 6, Application US/09632098
; Patent No. 6420154
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baird, Nand
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: MAMMALIAN ADHESION P
; FILE REFERENCE: 99-39
; CURRENT APPLICATION NUMBER: US/09/632,098
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ. ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 6
LENGTH: 2439
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2439)
; OTHER INFORMATION: n = A,T,C or G
US-09-632-098-6

```

```

OY      1825 GCACAGACAACATTGGACAGCTGCTGTCTCAAAACCATCTGTTCAGCAGCGCTCTGCAGAAATNCT 1884
        || .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      239 TCAAGTATGAATGAGATTTTCTTGATAACCCCATACCCGCTTTTTTAATCTTTAAGAAGCATCA 180
OY      1885 AATAATGACAGCGCCCTCAACGCTGGGGGTAGTTTAACCTCGAGACTTTTAGCCACTCTGGCA 1944
        || |..||..|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      179 AAATTTACACCCCATAGAGAAGCACGCCCTTTTTTTTAATFTGCATCATATTTCATCTACAAATGGCT 120
OY      1945 TCATATTCCTCCGTGCAACTTCTCAACCTGCTGCCCCCTGAGAGTCAACCAACTAT 1997
        || ..|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      119 TTACCTGATGATTCCTAAATTTGATTTGTGTCCTCTGTTACTGCACATGATAT 67

RESULT 10
US-09-117-257-22/c
; Sequence 22, Application US/09117257
; Patent No. 6214355
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; TITLE OF INVENTION: DDBP AND DBPB COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 4210..000500
; CURRENT APPLICATION NUMBER: US/09/117,257
; EARLIER FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: PCT/US96/17081
; EARLIER FILING DATE: 1996-10-22
; EARLIER APPLICATION NUMBER: 08/589,711
; EARLIER FILING DATE: 1996-01-22
; EARLIER APPLICATION NUMBER: 08/427,023
; EARLIER FILING DATE: 1995-04-24
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(561)
US-09-117-257-22

Query Match          1.3%: Score 34.6; DB 4; Length 561:
Best Local Similarity 46.8%; Pred. No. 0.5;
Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

OY      1765 GATTACACCTGGCAATGCTAGTACAGATTATATCCCTCTAGGAAAATTACATTAGGGCT 1824
        |||||...|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      299 GCTTCCTCTCTAATTTGATATAGCAAAATTTTCCGCTACTGTAGTAGCTGCGACATTTGCT 240
OY      1825 GCACACAGAACATTGGACAGCTGCTGTCTCAAAACCATCTGTTCAGCAGCGCTCTGCAGAAATNCT 1884
        |||||...|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      239 TCAAGTATGAATGAGATTTTCTTGATAACCCCATACCCGCTTTTTTAATCTTTAAGAAGCATCA 180
OY      1885 AATAATGACAGCGCCCTCAACGCTGGGGGTAGTTTAACCTCGAGACTTTTAGCCACTCTGGCA 1944
        || |..||..|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      179 AAATTTACACCCCATAGAGAAGCACGCCCTTTTTTTTAATFTGCATCATATTTCATCTACAAATGGCT 120
OY      1945 TCATATTCCTCCGTGCAACTTCTCAACCTGCTGCCCCCTGAGAGTCAACCAACTAT 1997
        || ..|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      119 TTACCTGATGATTCCTAAATTTGATTTGTGTCCTCTGTTACTGCACATGATAT 67

RESULT 11
US-09-117-257-24/c
; Sequence 24, Application US/09117257
; Patent No. 6214355
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; APPLICANT: Hanson, Mark
; TITLE OF INVENTION: DDBP AND DBPB COMPOSITIONS AND METHODS OF USE

```



Db 229 TCAAGTATGATGATTTCTGATACCCACTACCCGTTTTTTATCTTTAAAGCATCA 180  
QY 1885 AATAATGAGCGCCCTCAAGCTGGGGTGTAGTTACTCCGAGCTTTTACCATCTGGCA 1944  
Db 179 AAATTTACACCATAGAGAGCCCTTTTAAATGATCATCTATCTACAAATGGCT 120  
QY 1945 TCTATTTCCCTGCACACTTTCACACCTGCTGCCCTGAGAGTACCAACCTAT 1997  
Db 119 TTACGTGATGATTTCTAATTTGATTTTGTGCTCTCTGTAGTACCATGATAT 67

## RESULT 14

US-08-945-476-22/c  
Sequence 22, Application US/08945476  
Patent No. 6248517

## GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND  
METHODS OF USE  
NUMBER OF SEQUENCES: 27  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945,476  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/589,711  
FILING DATE: 22-JAN-1996  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/427,023  
FILING DATE: 24-APR-1995  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 561 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..561  
US-08-945-476-22

Query Match 1.3%; Score 34.6; DB 4; Length 561;  
Best Local Similarity 46.8%; Pred. No. 0.5;

Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1765 GATTACCTGCCAATGCTAGTACAGTTATATCTCTAGGAAATTTACATTAGGGCT 1824  
Db 299 GCTTCCTCTTCTATTTGCTATTACGAATTTTCCGCTACTGTAGTCTCGCACTTTTGT 240  
QY 1825 GCACCAAGAACATTTGACAGCTGCTTCAAAACCATCTGTAGAGCTCTCGAATACCT 1884  
Db 239 TCAAGTATGATGATTTTGTATGATACCCACACCCGTTTTTTATCTTTAAAGCATCA 180  
QY 1885 AATAATGAGCGCCCTCAAGCTGGGGTGTAGTTAACTCCGAGCTTTTACCATCTGGCA 1944  
Db 179 AAATTTACACCATAGAGAGCCCTTTTAAATGATCATCTATCTACAAATGGCT 120  
QY 1945 TCTATTTCCCTGCACACTTTCACACCTGCTGCCCTGAGAGTACCAACCTAT 1997  
Db 119 TTACGTGATGATTTCTAATTTGATTTTGTGCTCTCTGTAGTACCATGATAT 67

## RESULT 15

US-08-945-476-24/c  
Sequence 24, Application US/08945476  
Patent No. 6248517

## GENERAL INFORMATION:

TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND  
METHODS OF USE  
NUMBER OF SEQUENCES: 27  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945,476  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/589,711  
FILING DATE: 22-JAN-1996  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/427,023  
FILING DATE: 24-APR-1995  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 561 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..561  
US-08-945-476-24

Query Match 1.3%; Score 34.6; DB 4; Length 561;  
Best Local Similarity 46.8%; Pred. No. 0.5;

Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1765 GATTACCTGCCAATGCTAGTACAGTTATATCTCTAGGAAATTTACATTAGGGCT 1824  
Db 299 GCTTCCTCTTCTATTTGCTATTACGAATTTTCCGCTACTGTAGTCTCGCACTTTTGT 240  
QY 1825 GCACCAAGAACATTTGACAGCTGCTTCAAAACCATCTGTAGAGCTCTCGAATACCT 1884  
Db 239 TCAAGTATGATGATTTTGTATGATACCCACACCCGTTTTTTATCTTTAAAGCATCA 180  
QY 1885 AATAATGAGCGCCCTCAAGCTGGGGTGTAGTTAACTCCGAGCTTTTACCATCTGGCA 1944  
Db 179 AAATTTACACCATAGAGAGCCCTTTTAAATGATCATCTATCTACAAATGGCT 120  
QY 1945 TCTATTTCCCTGCACACTTTCACACCTGCTGCCCTGAGAGTACCAACCTAT 1997  
Db 119 TTACGTGATGATTTCTAATTTGATTTTGTGCTCTCTGTAGTACCATGATAT 67

Search completed: June 19, 2003, 04:27:58  
Job time : 138 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 23:11:38 : Search time 601 Seconds  
(without alignments)  
10139.613 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706  
Sequence: 1 atggcgtactatgaagcc.....cgctgacaggggaccttga 2706

Scoring table:  
IDENTITY-NUC  
Gap 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2706	100.0	2706	24	ABA95167
2	1896.8	70.1	4593	24	ABA95167
3	797.8	29.5	801	24	ABA95168
4	271.4	10.0	349	24	ABA95169
5	271.4	10.0	3715	24	ABA95170
6	113.2	4.2	438	24	ABU94207
7	67.6	2.5	865	21	AAC53156
8	47	1.7	4590	22	AAH24065
9	42.8	1.6	256	23	AAS58705

10	42.2	1.6	1595	24	AB199579
11	42.2	1.6	1893	22	AAH14157
12	41.6	1.5	2142	24	ABK63660
13	41.4	1.5	2052	24	AB199466
14	41	1.5	11680	21	AAC67600
15	39.6	1.4	6104	24	ABL33124
16	39	1.4	2190	21	AAZ51263
17	38.4	1.4	4668	22	AAH98410
18	38.4	1.4	5134	23	ABV24463
19	38	1.4	6161	23	ABL09562
20	37.8	1.4	4561	22	AAF32424
21	37.8	1.4	4907	22	AAFC3416
22	37.8	1.4	4946	11	AAO03115
23	37.8	1.4	5063	22	AAF32425
24	36.4	1.3	331	21	AAC01628
25	36.4	1.3	396	22	AAS37650
26	36.4	1.3	401	21	AAH30845
27	36.4	1.3	406	22	AAS37725
28	36.4	1.3	407	22	AAS37723
29	36.4	1.3	631	24	ABO56487
30	36.4	1.3	808	22	AAH05825
31	36.4	1.3	1501	22	AAH13847
32	36.4	1.3	3080	22	AAC91322
33	36.4	1.3	4069	24	ABN59810
34	36	1.3	4590	22	AAH24065
35	36	1.3	580073	18	AAT58840
36	35.8	1.3	6015	24	ABL32677
37	35.8	1.3	46340	21	ABN97978
38	35.6	1.3	1484	21	AAC47540
39	35.6	1.3	1486	21	AAC40456
40	35.6	1.3	2155561	24	ABN71557
41	35.4	1.3	3489	21	AAH30290
42	35.4	1.3	3489	22	AAH82901
43	35.4	1.3	3489	24	ABA93487
44	35.4	1.3	5059	20	AAH4332
45	35.4	1.3	32207	20	AAV73805

#### ALIGNMENTS

RESULT 1					
ID	ABA95167	standard; CDNA: 2706 BP.			
XX	ABA95167;				
AC	20-MAY-2002	(first entry)			
XX	20-MAY-2002				
DE	Arabidopsis floral induction gene (FPA) cDNA sequence.				
XX	FPA: floral induction; photoperiod; plant; flowering; FLC; FRI; frigidia;				
KW	flowering locus C; transgenic; gene; ss.				
XX	Arabidopsis thaliana.				
OS	Arabidopsis thaliana.				
XX	Key	Location/Qualifiers			
EH	1..2706				
FT	CDs				
FT	misc_binding	/tag= a			
FT	misc_binding	/product= "FPA protein"			
FT	misc_binding	/tag= b			
FT	misc_binding	/note= "RNA binding region"			
FT	misc_binding	/tag= c			
FT	misc_binding	/note= "RNA binding region"			
FT	misc_binding	/tag= d			
FT	misc_binding	/note= "RNA binding region"			
XX	W0200212518-A2.				
PN	14-FEB-2002.				
XX	14-FEB-2002.				
PD	14-FEB-2002.				

Mouse ischaemic co  
Human cDNA sequenc  
Rat sequence diffe  
Mouse ischaemic co  
Human ORF ORF2255  
Human immune syste  
Human RNA-associat  
Rat EST-derived co  
Human prostate exp  
Drosophila melanog  
L. mesenteroides s  
Nucleotide sequenc  
L. mesenteroides s  
Human secreted pro  
Novel human diagno  
Human colon cancer  
Novel human diagno  
Novel human diagno  
Human colon cancer  
Human cDNA clone (K  
Human cDNA sequenc  
Human polynucleoti  
Novel human coding  
Yeast AOD9604-asso  
Mycoplasma genital  
Human immune syste  
Human retroviral s  
Arabidopsis thalia  
Streptococcus poly  
Kaposi's sarcoma-a  
Nucleotide sequenc  
Kaposi's sarcoma-a  
Stealth virus nucl  
KSHV LTR DNA (nucl

XX 02-AUG-2001; 2001WO-US24427.  
 XX  
 XX PR 03-AUG-2000; 2000US-222550P.  
 XX  
 XX PA (MISC.) WISCONSIN ALUMNI RES FOUND.  
 XX  
 XX PA Amasino RM, Schomburg FM, Michaels SD, Patton D;  
 XX  
 XX PI WPI; 2002-227160/28.  
 XX  
 XX DR P-PSDB; ABB07659.  
 XX  
 XX  
 XX Novel isolated DNA sequence comprising coding sequence for floral  
 PT induction gene (FPA) gene which controls flowering time in plants,  
 PT useful for altering flowering time of plants and for downregulating  
 PT flowering locus C (FLC) mRNA activity  
 XX  
 XX  
 PS Claim 5; Page 29-34; 39pp; English.  
 XX  
 XX The invention relates to the FPA gene (floral induction promoter in  
 CC plants during both long and short day photoperiods) from Arabidopsis  
 CC thaliana. The FPA coding sequence is useful for altering flowering time  
 CC in a plant, and for down-regulating flowering locus C (FLC) mRNA  
 CC activity. The FPA gene can accelerate flowering in several late-flowering  
 CC mutant backgrounds and can fully compensate for addition of two naturally  
 CC occurring backgrounds, *frigida* (*FRI*) and *FLC* which confer late-flowering  
 CC phenotypes. Over expression of FPA can compensate for the delaying effect  
 CC caused by short days on floral induction, and decreases *FLC* mRNA in  
 CC plants containing the *FLC* gene. Fragments of the FPA gene can also act to  
 CC decrease activity of an endogenous FPA gene by modifying the expression  
 CC of the endogenous FPA gene and expression of a portion of polypeptide  
 CC -encoded by FPA gene can also lead to a delay of flowering in a plant. The  
 CC present sequence represents the A. thaliana FPA protein encoding cDNA.  
 XX  
 XX Sequence 2706 BP; 779 A; 628 C; 617 G; 682 T; 0 other;

Query Match	100.0%	Score 2706;	DB 24;	Length 2706;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2706; Conservative	0;	Mismatches	0;	Indels 0

QY	1	ATGGCGTTTTCATGAGCACTTCAGAGCCGATGATTCGGGTTTCCAGTCCAAACAATCTT	60
Dp	1	ATGGCGTTTTCATGAGCACTTCAGAGCCGATGATTCGGGTTTCCAGTCCAAACAATCTT	60
QY	61	TGGGTGGTAGCCTTAACGCCGAGACGACAGAGTCAGATCGACCGAGTCTTTGGAA	120
Dp	61	TGGGTGGTAGCCTTAACGCCGAGACGACAGAGTCAGATCGACCGAGTCTTTGGAA	120
QY	121	TACGGCGATTTGATAGAAATACGGGTGTAATCTTCAAGAGCTTTCGTTTATATAC	180
Dp	121	TACGGCGATTTGATAGAAATACGGGTGTAATCTTCAAGAGCTTTCGTTTATATAC	180
QY	181	AGACATGTGGAGAGCAGTCGACAGCCAAAGAGCTTTCAGAGACAAATTTGAAATGA	240
Dp	181	AGACATGTGGAGAGCAGTCGACAGCCAAAGAGCTTTCAGAGACAAATTTGAAATGA	240
QY	241	AGTCAAAATTAAGATGGAATACGACAGCCGGCAAAACCTTGTAAAGTCTATGGTGGT	300
Dp	241	AGTCAAAATTAAGATGGAATACGACAGCCGGCAAAACCTTGTAAAGTCTATGGTGGT	300
QY	301	GGAAATCGGCCCTAATGTCCTCAAGAGATGACCTGGAGAGAGTTCAGAAAGTTGGGAA	360
Dp	301	GGAAATCGGCCCTAATGTCCTCAAGAGATGACCTGGAGAGAGTTCAGAAAGTTGGGAA	360
QY	361	ATCGAGAGATTTAGGTTTCTCAGAGAACGCAAGACAGCTTTCATGATATTTATGAGATG	420
Dp	361	ATCGAGAGATTTAGGTTTCTCAGAGAACGCAAGACAGCTTTCATGATATTTATGAGATG	420
QY	421	GATGATGCTTTACAGGCTTAAGAGCATTAATGGAACCTATGGGTGAGCTTTTGGCT	480
Dp	421	GATGATGCTTTACAGGCTTAAGAGCATTAATGGAACCTATGGGTGAGCTTTTGGCT	480
QY	481	GTTGATTTTTCGCGGTACAGAGCCCAAAAAAGAACCAATGGGCTGGCTCTTACGATAC	540

[illegible]

Db	1501	CGGGCGGGGTGTTCCCAATTAGATGATGATGATCAACTTTATTCTTGTCGGCCCTCATCAEAT	1620
Qy	1621	TTCTTAACGTATGTACTCCAAAGTAGACCCGTCACAGAGGGCTATATGGTGTGTCTCAAG	1680
Db	1621	TTCTTAACGTATGTACTCCAAAGTAGACCCGTCACAGAGGGCTATATGGTGTGTCTCAAG	1680
Qy	1681	TTACCCCGCGCAGCCGTTCTGTGTACAGCATCATACAGACAAGATCTCACTCCAATCT	1740
Db	1681	TTACCCCGCGCAGCCGTTCTGTGTACAGCATCATACAGACAAGATCTCACTCCAATCT	1740
Qy	1741	CTGCATTATATGATGATCAAGCCCGGGATTTACCTGCGCAATGCTACTGCACAGTTTATTCCT	1800
Db	1741	CTGCATTATATGATGATCAAGCCCGGGATTTACCTGCGCAATGCTACTGCACAGTTTATTCCT	1800
Qy	1801	CCTAGGGGAAATTACATTAGGGGTGACACCAACACATTGTGACAGCTGCTTCAAAACCATCT	1860
Db	1801	CCTAGGGGAAATTACATTAGGGGTGACACCAACACATTGTGACAGCTGCTTCAAAACCATCT	1860
Qy	1861	GTTAGCCAGCCTCTCAGATATCTTAATATATGACGCGCTCAAGCTGGGGTTAGTTTAAT	1920
Db	1861	GTTAGCCAGCCTCTCAGATATCTTAATATATGACGCGCTCAAGCTGGGGTTAGTTTAAT	1920
Qy	1921	CCGGAGCTTTTAGCACACTGTGGCAATCTATCTCCCTGCAATCTGTCAACCTGCGCCCT	1980
Db	1921	CCGGAGCTTTTAGCACACTGTGGCAATCTATCTCCCTGCAATCTGTCAACCTGCGCCCT	1980
Qy	1981	GAGAGTCACCAACCTATGTACAGACCTTCACACATGTTGTTCCACAGCATCAGTCCAAAT	2040
Db	1981	GAGAGTCACCAACCTATGTACAGACCTTCACACATGTTGTTCCACAGCATCAGTCCAAAT	2040
Qy	2041	GGACTGTACAAATGAGAGAAGCACCGCTTCACAGCTTGGAAAAGAGGTCCCAAAACAGTTTAT	2100
Db	2041	GGACTGTACAAATGAGAGAAGCACCGCTTCACAGCTTGGAAAAGAGGTCCCAAAACAGTTTAT	2100
Qy	2101	GATCGCTCAAAATCAATGATATTCACACAAATAGGGAATAGTCACTCCACAGCTGGGCACTA	2160
Db	2101	GATCGCTCAAAATCAATGATATTCACACAAATAGGGAATAGTCACTCCACAGCTGGGCACTA	2160
Qy	2161	CCTCCTCCTCCTTTCGCGCTTTACCTCCACAGCTTCAAAACAACCCCACTACACTAGTGGAAATG	2220
Db	2161	CCTCCTCCTCCTTTCGCGCTTTACCTCCACAGCTTCAAAACAACCCCACTACACTAGTGGAAATG	2220
Qy	2221	GTCATATGGCAACATGCAATTACCAAGGCCAATCTGTAAACAGCTCCAGCTGTCCGCTTA	2280
Db	2221	GTCATATGGCAACATGCAATTACCAAGGCCAATCTGTAAACAGCTCCAGCTGTCCGCTTA	2280
Qy	2281	CCAAATATGCTCCATTAATTAATTTATTCATGTACACTAGGGTTGGTCAAAATCAATCTGTT	2340
Db	2281	CCAAATATGCTCCATTAATTAATTTATTCATGTACACTAGGGTTGGTCAAAATCAATCTGTT	2340
Qy	2341	TCTCAGCCCATGTATCCAGCAATTACCAACCAACAGCCGTCACATGCCAAACCAAACTATAGT	2400
Db	2341	TCTCAGCCCATGTATCCAGCAATTACCAACCAACAGCCGTCACATGCCAAACCAAACTATAGT	2400
Qy	2401	CCAATTCAGATATATCAGCAAGCTAATTTTCATGGCGTAACCAACAAATCAGGCACAGAAC	2460
Db	2401	CCAATTCAGATATATCAGCAAGCTAATTTTCATGGCGTAACCAACAAATCAGGCACAGAAC	2460
Qy	2461	TTTAAACCTTCCCAATTTTCAAGCGTCGATACCAACACACAGAGATTAAGGCCAAATTTAGAG	2520
Db	2461	TTTAAACCTTCCCAATTTTCAAGCGTCGATACCAACACACAGAGATTAAGGCCAAATTTAGAG	2520
Qy	2521	CCACAAAAACCAAGACTATCAGATTTGACGCTTATGATCTCTGGGGATGGTCAAGGGTACACA	2580
Db	2521	CCACAAAAACCAAGACTATCAGATTTGACGCTTATGATCTCTGGGGATGGTCAAGGGTACACA	2580
Qy	2581	GATGGGGAGTGCATTAAGAAATCAGAGATTACCAAGTCAACACTTCAATTTGCAGCAAACTT	2640
Db	2581	GATGGGGAGTGCATTAAGAAATCAGAGATTACCAAGTCAACACTTCAATTTGCAGCAAACTT	2640
Qy	2641	CTTCTCCAGATACAGCAGAAACACAGCAGACAGCTTTCAGTACTCCGGCTGTGACAGGGC	2700
Db	2641	CTTCTCCAGATACAGCAGAAACACAGCAGACAGCTTTCAGTACTCCGGCTGTGACAGGGC	2700

QY	2701	CCTGA 2706	
Db	2701	CCTGA 2706	
RESULT 2			
ID	ABA95166	standard; DNA: 4593 BP.	
XX	ABA95166;		
AC			
XX			
XX	20-MAY-2002	(first entry)	
DE			
XX	Arabidopsis floral induction gene (FPA) sequence.		
XX			
KW	FPA; floral induction; photoperiod; plant; flowering; FLC; FRI; frigida;		
KW	flowering locus C; transgenic; gene; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
FN	WO200212518-A2.		
PD			
XX	14-FEB-2002.		
PF			
XX	02-AUG-2001; 2001WO-US24427.		
FR			
XX	03-AUG-2000; 2000US-222550P.		
PA	(WISC ) WISCONSIN ALUMNI RES FOUND.		
XX			
PI	Amasino RM, Schomburg FM, Michels SD, Patton D;		
DR	WPI: 2002-227160/28.		
XX	P-PSDB: ABB07659.		
PT	Novel isolated DNA sequence comprising coding sequence for floral		
PT	induction gene (FPA) gene which controls flowering time in plants,		
PT	useful for altering flowering time of plants and for downregulating		
XX	flowering locus C (FLC) mRNA activity -		
PS	Claim 5; Page 27-29; 39p; English.		
CC	The invention relates to the FPA gene (floral induction promoter in		
CC	plants during both long and short day photoperiods) from Arabidopsis		
CC	thaliana. The FPA coding sequence is useful for altering flowering time		
CC	in a plant, and for down-regulating flowering locus C (FLC) mRNA		
CC	activity. The FPA gene can accelerate flowering in several late-flowering		
CC	mutant backgrounds and can fully compensate for addition of two naturally		
CC	occurring genes, frigida (FRI) and FLC which confer late-flowering		
CC	phenotypes. Over expression of FPA can compensate for the delaying effect		
CC	caused by short days on floral induction, and decreases FLC mRNA in		
CC	plants containing the FLC gene. Fragments of the FPA gene can also act to		
CC	decrease activity of an endogenous FPA gene by modifying the expression		
CC	of the endogenous FPA gene and expression of a portion of polypeptide		
CC	encoded by FPA gene can also lead to a delay of flowering in a plant.		
CC	The present sequence represents the A. thaliana FPA gene sequence.		
SO	Sequence 4593 BP; 1275 A; 958 C; 978 G; 1382 T; 0 other;		
QY	Query Match	70.1%; Score 1896.8; DB 24; Length 4593;	
Db	Best Local Similarity	90.2%; Pred. No. 0;	
	Matches 2137; Conservative	0; Mismatches 2; Indels 229; Gaps 1;	
QY	568	CAGTATCTCCTCACTCATATGAAGACTTTAAAGAGATGTCAGCCAGTAAGCTTGTGG 627	
Db	2226	CAGTATCCTCCTCACTCATATGAAGACTTTAAAGAGATGTCAGCCAGTAAGCTTGTGG 2285	
QY	628	ATTGGGTTCCCTCCTCACTCACTGTCACCAATGCAATGATGCAAAATTCGACAAATGCATG 687	
Db	2286	ATTGGGTTCCCTCCTCACTCACTGTCACCAATGCAATGATGAGCAAAATTCGACAAATGCATG 2345	
QY	688	ATTACTCTTTGGTGTGAGATCGAAGGGTAAAGAAAGTTACCATCAGCAATTTTGGCACTTGTG 747	

Db 2346 ATACTCTTGGATGAGATGAGAGGGTAAAGTTTCCATCAAGAAATTTTGCACTTGTG 2405  
 QY 748 GAGTTAGAGCGCGGAGAAAGCTCCCAATGCAAGAAAGCCTACAGGGAGGTTATTC 807  
 Db 2406 GAGTTAGAGCGCGGAGAAAGCTCCCAATGCAAGAAAGCCTACAGGGAGGTTATTC 2465  
 QY 808 AATAATCTAGAAATTAATGATGCAAAAGCAAGTGTGCTCCCTGACACAGAGAT 867  
 Db 2466 AATAATCTAGAAATTAATGATGCAAAAGCAAGTGTGCTCCCTGACACAGAGAT 2525  
 QY 868 ACTAGTTTACTCTGTATGAAACGCTCAAGAGACAGATATGTTCAATATGATCTTCA 927  
 Db 2526 ACTAGTTTACTCTGTATGAAACGCTCAAGAGACAGATATGTTCAATATGATCTTCA 2585  
 QY 928 TGTGATGCTCTCCATCTTACGTAATCTCTGGGCTATAGAGCCCTCAGAGTACG 987  
 Db 2586 TTTGATGCTCTCCATCTTACGTAATCTCTGGGCTATAGAGCCCTCAGAGTACG 2645  
 QY 988 AATGAGCGTTCATATATGATGCAAGATACAAATGACGTTGTTGTAAGAGCCAAAGTGG 1047  
 Db 2646 AATGAGCGTTCATATATGATGCAAGATACAAATGACGTTGTTGTAAGAGCCAAAGTGG 2705  
 QY 1048 AGGAGGCGATCTGCAAAATGGAACCTGGAATACCTCCATCTCCACAGAGACCTGGAATCTC 1107  
 Db 2706 AGGAGGCGATCTGCAAAATGGAACCTGGAATACCTCCATCTCCACAGAGACCTGGAATCTC 2765  
 QY 1108 CCATCTCTCTGCAAGAGTACGAGGCGCCTATGAGTCAAAACCCGATCTCTGGAGAGA 1167  
 Db 2766 CCATCTCTCTGCAAGAGTACGAGGCGCCTATGAGTCAAAACCCGATCTCTGGAGAGA 2825  
 QY 1168 TATATCTCTGCTCAAGTTGGTCAGAGAAATGAACGAACGAAAGAGATGATCAGTGAC 1227  
 Db 2826 TATATCTCTGCTCAAGTTGGTCAGAGAAATGAACGAACGAAAGAGATGATCAGTGAC 2885  
 QY 1228 GGTTCCTCTCAATGGGTGTCGATGAGAGGTCATTTGTCAGGTTCAAGTTGCTGCTAGA 1287  
 Db 2886 GGTTCCTCTCAATGGGTGTCGATGAGAGGTCATTTGTCAGGTTCAAGTTGCTGCTAGA 2945  
 QY 1288 CCAATCCGTGGCCCCCTGATTTCTGATCACATATGAGAGAAATGATTCGCAAGGGTGA 1347  
 Db 2946 CCAATCCGTGGCCCCCTGATTTCTGATCACATATGAGAGAAATGATTCGCAAGGGTGA 3005  
 QY 1348 ACTCCGCTGTGTGCTGCTGTCATACCATATGAGAAAGGGATTGAACTAACT-- 1403  
 Db 3006 ACTCCGCTGTGTGCTGCTGTCATACCATATGAGAAAGGGATTGAACTAACTGCA 3065  
 QY 1404 ----- 1403  
 Db 3066 GTACTAATTTCTAGCACTTAAACCTTCTAGTGTTCCTTTTTCAGAGGATTTATATAT 3125  
 QY 1404 ----- 1403  
 Db 3126 TTTCCATTTTCATTCGATGAGAGTAACATATATATAGATAGTATTTTATTTACTA 3185  
 QY 1404 ----- 1403  
 Db 3186 TTAATGTTTGTAGTTTGTAGATGCTTGAATTTTCAATGCTGTTGATTCATTTTGGCATTG 3245  
 QY 1404 ----- 1403  
 Db 3246 CCCTCAATTAAGTACTGTTTGTCTTTTATTAATGATTTATAGCCCTGAGGTGCTCA 3305  
 QY 1419 TTGTTAGCAAGAACTGATTTGAATATGCTGCTAAACCTTACGCCGTTGCCATTGATG 1478  
 Db 3306 TTGTTAGCAAGAACTGATTTGAATATGCTGCTAAACCTTACGCCGTTGCCATTGATG 3365  
 QY 1479 TGAGATCGTTTTTTCGTACCAAGAGAGAAAGATTTTGGCTCTTACACTGAATTTCT 1538  
 Db 3366 TGAGATCGTTTTTTCGTACCAAGAGAGAAAGATTTTGGCTCTTACACTGAATTTCT 3425  
 QY 1539 CCGGTACTTACCTCAAAAGATCGGGGCTTCCCAATATAGATGATGATCAACTTT 1598

Db 3426 CCGGTACTTACCTCAAAAGATCGGGGCTTCCCAATATAGATGATGATCAACTTT 3485  
 QY 1599 ATTCTGTGCTCCTCATCAGATTTCTTAACGATGATCTCAAGTACGACCCGTAAGAAC 1658  
 Db 3486 ATTCTGTGCTCCTCATCAGATTTCTTAACGATGATCTCAAGTACGACCCGTAAGAAC 3545  
 QY 1659 GCTATATGTTGTTCTTCAAGTATACCCCGCAGCCGCTCTGTTACAGCATATACAG 1718  
 Db 3546 GCTATATGTTGTTCTTCAAGTATACCCCGCAGCCGCTCTGTTACAGCATATACAG 3605  
 QY 1719 ACAAGATCTCAGTCCCATCTCTCGATTTATGATCAAGCCCGGATTCACCTGCCAA 1778  
 Db 3606 ACAAGATCTCAGTCCCATCTCTCGATTTATGATCAAGCCCGGATTCACCTGCCAA 3665  
 QY 1779 TGTAGTACAGTTTATATCTCTCTAGGAAATTCATATAGGGGTGACACGAACATTT 1838  
 Db 3666 TGTAGTACAGTTTATATCTCTCTAGGAAATTCATATAGGGGTGACACGAACATTT 3725  
 QY 1839 GACAGCTGCTTCAAAACCATCTGTATGCGAGCCTTCAGAAATCTTAATATGACAGCC 1898  
 Db 3726 GACAGCTGCTTCAAAACCATCTGTATGCGAGCCTTCAGAAATCTTAATATGACAGCC 3785  
 QY 1899 TCAAGCTGGGTTAGTTTAACTCCGAGGCTTTAGCCAGTCCGATCTATCTCCGTC 1958  
 Db 3786 TCAAGCTGGGTTAGTTTAACTCCGAGGCTTTAGCCAGTCCGATCTATCTCCGTC 3845  
 QY 1959 AACTTTCAACTGCTGCTGCTGAGTACCAACCTATGTCAGACCTTCAACAGTTGT 2018  
 Db 3846 AACTTTCAACTGCTGCTGCTGAGTACCAACCTATGTCAGACCTTCAACAGTTGT 3905  
 QY 2019 TTCCACAGACATCACTCCCAATGAGTATGACATGAGAGAAAGCCGTCACAGCTTGAA 2078  
 Db 3906 TTCCACAGACATCACTCCCAATGAGTATGACATGAGAGAAAGCCGTCACAGCTTGAA 3965  
 QY 2079 AAGAGTCCACAAACAGTTTCATGATGCTCAATCAATCACTATTCACAAATAGGAATCA 2138  
 Db 3966 AAGAGTCCACAAACAGTTTCATGATGCTCAATCAATCACTATTCACAAATAGGAATCA 4025  
 QY 2139 GTACACTGCAAGTGGGCACTACCTCTCTCTGCTGCTGCTTACCTCCAGCTTCAACAA 2198  
 Db 4026 GTACACTGCAAGTGGGCACTACCTCTCTCTGCTGCTGCTTACCTCCAGCTTCAACAA 4085  
 QY 2199 CCCCAACTACATAGTGAATGTCATAGGCAACATGCAATTCAGAGCAATCTGTAA 2258  
 Db 4086 CCCCAACTACATAGTGAATGTCATAGGCAACATGCAATTCAGAGCAATCTGTAA 4145  
 QY 2259 CATGCTCAGCTGCTCCGTTACCAATATGCTCTATATATTTTCCATGATCACTCA 2318  
 Db 4146 CATGCTCAGCTGCTCCGTTACCAATATGCTCTATATATTTTCCATGATCACTCA 4205  
 QY 2319 GGGTTCGTCAAATCACTGTTCTTCAGCCCATGCTCGAGAAATTCACAGCAAGCGTC 2378  
 Db 4206 GGGTTCGTCAAATCACTGTTCTTCAGCCCATGCTCGAGCAATTCACAGCAAGCGTC 4265  
 QY 2379 CATGCCAAACCAAACTATGTCATTCGAATTCAGAGTTATCAGCAAGCTAATTTTCAATGGCGT 2438  
 Db 4266 CATGCCAAACCAAACTATGTCATTCGAATTCGAATTCAGCAAGCTAATTTTCAATGGCGT 4325  
 QY 2439 AACACCAATCAGGACAGAACTTAACCTTCCCAATTTCAACCTGCAATGCAACCAAC 2498  
 Db 4326 AACACCAATCAGGACAGAACTTAACCTTCCCAATTTCAACCTGCAATGCAACCAAC 4385  
 QY 2499 AGCAGATTAAGCAAAATTTAGAGCCCAAAACCAAGCACTAGATGAGCTTATGATCTC 2558  
 Db 4386 AGCAGATTAAGCAAAATTTAGAGCCCAAAACCAAGCACTAGATGAGCTTATGATCTC 4445  
 QY 2559 TGGGATGCTCAGGTTACAAAGATGGGAGGTGATTAAGATCAGAGATACAGTCAAC 2618  
 Db 4446 TGGGATGCTCAGGTTACAAAGATGGGAGGTGATTAAGATCAGAGATACAGTCAAC 4505  
 QY 2619 ACTACAAATTTGACAGAAACCTTCTTCACATACAGAGAAAGAGAGCAAGAGCTTTC 2678  
 Db 4506 ACTACAAATTTGACAGAAACCTTCTTTCACATACAGAGAAAGAGAGCAAGAGCTTTC 4565

OY 2679 AGGTACTCCGGCTGGACAGGGGCTTGA 2706  
DB 4566 AGGTACTCCGGCTGGACAGGGGCTTGA 4593

RESULT 3  
ABA95168/c  
ID ABA95168 standard; DNA; 801 BP.  
XX ABA95168;  
AC  
XX 20-MAY-2002 (first entry)  
DE Arabidopsis floral induction gene (FPA) antisense fragment.  
XX FPA; floral induction; photoperiod; plant; flowering; FLC; FRI; frigida;  
KW flowering locus C; transgenic; gene; antisense; ds.  
XX Arabidopsis thaliana.  
OS  
XX WO200212518-A2.  
PN  
XX 14-FEB-2002.  
PD  
XX 02-AUG-2001; 2001WO-US24427.  
PE  
XX 03-AUG-2000; 2000US-22250P.  
PR  
XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
PA  
XX Amasino RM, Schomburg FM, Michaels SD, Patton D;  
PI WPI: 2002-227160/28.  
DR  
XX Novel isolated DNA sequence comprising coding sequence for floral  
PT induction gene (FPA) gene which controls flowering time in plants,  
PT -useful for altering flowering time of plants and for downregulating  
PT flowering locus C (FLC) mRNA activity -  
XX  
XX Claim 5; Page 37; 39pp; English.  
PS  
XX The invention relates to the FPA gene (floral induction promoter in  
CC plants during both long and short day photoperiods) from Arabidopsis  
CC thaliana. The FPA coding sequence is useful for altering flowering time  
CC in a plant, and for down-regulating flowering locus C (FLC) mRNA  
CC activity. The FPA gene can accelerate flowering in several late-flowering  
CC mutant backgrounds and can fully compensate for addition of two naturally  
CC occurring genes, frigida (FRI) and FLC which confer late-flowering  
CC phenotypes. Over expression of FPA can compensate for the delaying effect  
CC caused by short days on floral induction, and decreases FLC mRNA in  
CC plants containing the FLC gene. Fragments of the FPA gene can also act to  
CC decrease activity of an endogenous FPA gene by modifying the expression  
CC of the endogenous FPA gene and expression of a portion of polypeptide  
CC encoded by FPA gene can also lead to a delay of flowering in a plant. The  
CC present sequence represents the A. thaliana FPA gene antisense fragment.  
XX  
XX Sequence 801 BP; 206 A; 207 C; 166 G; 222 T; 0 other;  
SQ

Query Match 29.5%; Score 797.8; DB 24; Length 801;  
Best Local Similarity 99.8%; Pred. No. 1.6e-240;  
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 587 AAGACTTTAAGAGAGATGTCAGCCAAAGTAAGTGTGTGTGGATTGGCTTCCCTCTACTG 646  
DB 801 AAGACTTTAAGAGAGATGTCAGCCAAAGTAAGTGTGTGTGGATTGGCTTCCCTCTACTG 742  
OY 647 CTACACATGCAATGAGAGCAATTCCTGCACATGCGATATCTCTTTGCTGATGCG 706  
DB 741 CTACACATGCAATGAGAGCAATTCCTGCACATGCGATATCTCTTTGCTGATGCG 682  
OY 707 AGAGGTTAAAGTTACCATCAAGAAATTTGCACTGTGAGTTTGGAGCCGCGAGG 766  
|||||

DB 681 AGAGGTTAAAGTTACCATCAAGAAATTTGCACTGTGAGTTTGGAGCGCGAGG 622  
OY 767 AAGCTCGCAATGCAAGAGAGGCTTACAGGGGAGTTATTCAATATCTTGAAATCAAAA 826  
DB 621 AAGCTCGCAATGCAAGAGAGGCTTACAGGGGAGTTATTCAATATCTTGAAATCAAAA 562  
OY 827 TTATGTACTCAAGAGATGAGTTGCTCTCTGAGCAAGAGATAGTATTGTTTACTCTGTA 886  
DB 561 TTATGTACTCAAGAGATGAGTTGCTCTCTGAGCAAGAGATAGTATTGTTTACTCTGTA 502  
OY 887 TGAACGCTCAAGAGAGATATGTCATATATGATTCCTTCAATGATATCTTCTCTCAT 946  
DB 501 TGAACGCTCAAGAGAGATATGTCATATATGATTCCTTCAATGATATCTTCTCTCAT 442  
OY 947 CTACTGCAATTCCTGCTGCTATGAGGCCCTCAGAGAGTACAGATAGAGCTTCATATATG 1006  
DB 441 CTACTGCAATTCCTGCTGCTATGAGGCCCTCAGAGAGTACAGATAGAGCTTCATATATG 382  
OY 1007 GTGCGAATACAAAGAGCTGTGTGTAAGAGCCAACTGGAGAGGCCATCTGCAATG 1066  
DB 381 GTGCGAATACAAAGAGCTGTGTGTAAGAGCCAACTGGAGAGGCCATCTGCAATG 322  
OY 1067 GAAGTGAATATCTCCATCTCCAAAGAGAGCTGATCTCCATCTCTCCACAAAGTA 1126  
DB 321 GAAGTGAATATCTCCATCTCCAAAGAGAGCTGATCTCCATCTCTCCACAAAGTA 262  
OY 1127 CGAGGCGCCCTATGAGTCAAAACCCCATTTCTTGGGAAGATATGATCTCTGCTCATG 1186  
DB 261 CGAGGCGCCCTATGAGTCAAAACCCCATTTCTTGGGAAGATATGATCTCTGCTCATG 202  
OY 1187 TCAGAGAAATGAAGCAACGAGAGATGATGATGATGATGATGATGATGATGATGATG 1246  
DB 201 TCAGAGAAATGAAGCAACGAGAGATGATGATGATGATGATGATGATGATGATGATG 142  
OY 1247 TCGATGAGAGTCAATTTGCTGAGGTTCACTGCTGCTAGACTATCCGTGCGCCCTG 1306  
DB 141 TCGATGAGAGTCAATTTGCTGAGGTTCACTGCTGCTAGACTATCCGTGCGCCCTG 82  
OY 1307 ATTCTGATCACAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1366  
DB 81 ATTCTGATCACAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 22

OY 1367 GTTGTACTCATGGGAAGG 1387  
DB 21 GTTGTACTCATGGGAAGG 1

RESULT 4  
ABA95169  
ID ABA95169 standard; DNA; 349 BP.  
XX ABA95169;  
AC  
XX 20-MAY-2002 (first entry)  
DE Arabidopsis floral induction gene (FPA) coding fragment.  
XX FPA; floral induction; photoperiod; plant; flowering; FLC; FRI; frigida;  
KW flowering locus C; transgenic; gene; ds.  
XX Arabidopsis thaliana.  
OS  
XX WO200212518-A2.  
PN  
XX 14-FEB-2002.  
PD  
XX 02-AUG-2001; 2001WO-US24427.  
PE  
XX 03-AUG-2000; 2000US-22250P.  
PR  
XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
PA  
XX Amasino RM, Schomburg FM, Michaels SD, Patton D;  
PI

XX WPI: 2002-227160/28.

DR Novel isolated DNA sequence comprising coding sequence for floral  
XX induction gene (FPA) gene which controls flowering time in plants,  
PT useful for altering flowering time of plants and for downregulating  
PT flowering locus C (FLC) mRNA activity

XX Claim 5; Page 37; 39pp; English.

XX The invention relates to the FPA gene (floral induction promoter in  
CC plants during both long and short day photoperiods) from Arabidopsis  
CC thaliana. The FPA coding sequence is useful for altering flowering time  
CC in a plant, and for down-regulating flowering locus C (FLC) mRNA  
CC activity. The FPA gene can accelerate flowering in several late-flowering  
CC mutant backgrounds and can fully compensate for addition of two naturally  
CC occurring genes, *frigida* (FRI) and *FLC* which confer late-flowering  
CC phenotypes. Over expression of FPA can compensate for the delaying effect  
CC caused by short days on floral induction, and decreases FLC mRNA in  
CC plants containing the FLC gene. Fragments of the FPA gene can also act to  
CC decrease activity of an endogenous FPA gene by modifying the expression  
CC of the endogenous FPA gene and expression of a portion of polypeptide  
CC encoded by FPA gene can also lead to a delay of flowering in a plant. The  
CC present sequence represents the A. thaliana FPA gene coding fragment.

XX Sequence 349 BP; 92 A; 68 C; 81 G; 108 T; 0 other;

Query Match 10.0%; Score 271.4; DB 24; Length 349;

Best Local Similarity 94.6%; Pred. No. 1,1e-74;

Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 60  
DB 1 ATGGCGTTATCTATGAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 60  
QY 61 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGGAGA 120  
DB 61 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGGAGA 120  
QY 121 TACGGCGATATGATAGCAATCAGCGTGTATCTTCAAGAGCTTTGGTTATATCTAC 180  
DB 121 TACGGCGATATGATAGCAATCAGCGTGTATCTTCAAGAGCTTTGGTTATATCTAC 180  
QY 181 AGCAGTGTGAG 240  
DB 181 AGCAGTGTGAG 240  
QY 241 AGTCAATTTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297  
DB 241 AGTCAATTTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297

RESULT 5

ABA95170 standard; DNA; 3715 BP.

AC ABA95170;

DT 20-MAY-2002 (first entry)

DE Arabidopsis floral induction gene (FPA) fragment.

KW FPA; floral induction; photoperiod; plant; flowering; FLC; FRI; *frigida*;  
KM flowering locus C; transgenic; gene; ds.

OS Arabidopsis thaliana.

XX MO200212518-A2.

XX 14-FEB-2002.

PF 02-AUG-2001; 2001WO-US24427.

XX

PR 03-AUG-2000; 2000US-222550P.

XX (WISC ) WISCONSIN ALUMNI RES FOUND.

XX Anasino RM, Schomburg FM, Michaels SD, Patton D;

DR WPI: 2002-227160/28.

PT Novel isolated DNA sequence comprising coding sequence for floral  
PT induction gene (FPA) gene which controls flowering time in plants,  
PT useful for altering flowering time of plants and for downregulating  
PT flowering locus C (FLC) mRNA activity

XX Claim 5; Page 38; 39pp; English.

XX The invention relates to the FPA gene (floral induction promoter in  
CC plants during both long and short day photoperiods) from Arabidopsis  
CC thaliana. The FPA coding sequence is useful for altering flowering time  
CC in a plant, and for down-regulating flowering locus C (FLC) mRNA  
CC activity. The FPA gene can accelerate flowering in several late-flowering  
CC mutant backgrounds and can fully compensate for addition of two naturally  
CC occurring genes, *frigida* (FRI) and *FLC* which confer late-flowering  
CC phenotypes. Over expression of FPA can compensate for the delaying effect  
CC caused by short days on floral induction, and decreases FLC mRNA in  
CC plants containing the FLC gene. Fragments of the FPA gene can also act to  
CC decrease activity of an endogenous FPA gene by modifying the expression  
CC of the endogenous FPA gene and expression of a portion of polypeptide  
CC encoded by FPA gene can also lead to a delay of flowering in a plant. The  
CC present sequence represents the A. thaliana FPA gene fragment comprising  
CC the FPA promoter and intron.

XX Sequence 3715 BP; 1055 A; 728 C; 711 G; 1221 T; 0 other;

Query Match 10.0%; Score 271.4; DB 24; Length 3715;

Best Local Similarity 94.6%; Pred. No. 4.9e-74;

Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 60  
DB 1832 ATGGCGTTATCTATGAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 1891  
QY 61 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGGAGA 120  
DB 1892 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGGAGA 1951  
QY 121 TACGGCGATATGATAGCAATCAGCGTGTATCTTCAAGAGCTTTGGTTATATCTAC 180  
DB 1952 TACGGCGATATGATAGCAATCAGCGTGTATCTTCAAGAGCTTTGGTTATATCTAC 2011  
QY 181 AGCAGTGTGAG 240  
DB 2012 AGCAGTGTGAG 2071  
QY 241 AGTCAATTTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297  
DB 2072 AGTCAATTTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2128

RESULT 6

ABL94207 standard; cDNA; 438 BP.

AC ABL94207;

DT 10-JUN-2002 (first entry)

DE Arabidopsis thaliana nucleic acid sequence Ref:2027972 SEQ ID NO:972.

KW Arabidopsis thaliana; insecticide; fungicide; plant; mapping; diagnosis;

XX genetic modification; gene; ss.

OS Arabidopsis thaliana.

XX

PN US2002023280-A1.  
XX 21-FEB-2002.  
XX 26-JAN-2001; 2001US-0770444.  
XX 27-JAN-2000; 2000US-178502P.  
XX (GORL) GORLACH J.  
PA (HANY) AN Y.  
PA (HAMI) HAMILTON C M.  
PA (PRIC) PRICE J L.  
PA (RAIN) RAINES T M.  
PA (YUYU) YU Y.  
PA (RAME) RAMEKA J G.  
PA (PAGE) PAGE A.  
PA (MATH) MATHEN A V.  
PA (LEDF) LEDFORD B L.  
PA (WOES) WOESSNER J P.  
PA (HAAS) HAAS W D.  
PA (GARC) GARCIA C A.  
PA (KRIC) KRICKER M.  
PA (SLAT) SLATER T.  
PA (DAVI) DAVIS K R.  
PA (ALLE) ALLEN K.  
PA (HOEF) HOFFMAN N.  
PA (HURB) HURBAN P.  
XX (GORLACH J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX WPI: 2002-267486/31.  
XX New Arabidopsis thaliana nucleic acid, for identifying homologous  
PT genes, producing compositions that modulate the expression or function  
PT of its encoded protein, and mapping functional regions of a protein -  
XX  
PS Claim 1; SEQ ID 972; 44pp; English.  
XX  
XX The present invention describes an Arabidopsis thaliana nucleic acid (1)  
CC comprising a sequence capable of hybridizing under stringent conditions  
CC to a sequence (S1) selected from any one of the 999 sequences given in  
CC AB193236 to AB194234. (1) have insecticide and fungicide activities, and  
CC they can be used as protein expression modulators. (1) can be used in  
CC identifying homologous or related genes, in producing compositions that  
CC modulate the expression or function of their encoded proteins, mapping  
CC functional regions of the proteins, and in studying associated  
CC physiological pathways. (1) can also be used: (1) for the genetic  
CC manipulation of cells, particularly plant cells; (2) in screening assays  
CC of various plant strains to determine the strains that are best capable  
CC of withstanding a particular disease or environmental stress; (3) for  
CC enhancing or inhibiting production of a biosynthetic product in a plant;  
CC (4) as probes in mapping and in diagnosis, in genetic modification and  
CC for screening purposes, to generate additional copies of the nucleic  
CC acids, to generate ribozymes or antisense oligonucleotides, and as  
CC single-stranded DNA probes or as triple-strand forming oligonucleotides;  
CC and (5) for generating genetically modified transgenic organisms.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site.  
XX  
XX Sequence 438 BP; 113 A; 79 C; 112 G; 134 T; 0 other;  
SQ

Query Match 4.2%; Score 113.2; DB 24; Length 438;  
Best Local Similarity 61.6%; Pred. No. 9.9e-25;  
Matches 181; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

OY 1390 ATGGAACCTAAACGCTGAGCGTCTCAATTGTCACAGAACTGATTGGAATATGCTC 1449  
DB 1 ATGACACATGATGCTGCTGCTGCTTCTAGATTGACAGCAAGAACTGTTTACATGCTG 60

OY 1450 GCTAAACATACGCCGCTTGCATTCATGATGATCGTTTTTTCGTACACAGACAGGAA 1509  
DB 61 GCTAAGCATTTACTACCAATCATCTTAAGACATGCGTGTTCCTTCTGCGAAGTAT 120  
OY 1510 GAAGATTTTGGCTCTTACACTGAATTTCTCCGTCACCTTAGCTCAAAAGATCGCGGCT 1569  
DB 121 GCTGATATCGTGTATTATGATGAGGTTTATGATATATCTGAGAGACAGACAGCGGACGT 180  
OY 1570 GTTGCCAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629  
DB 181 GTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
OY 1630 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1683  
DB 241 AAGATCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 294

RESULT 7  
AAC53156  
ID AAC53156 standard; DNA; 865 BP.  
XX AAC53156;  
AC 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 73564.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS  
XX  
XX EPI033405-A2.  
XX  
XX 06-SEP-2000.  
PD  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 07-MAY-1999; 99US-0132487.  
XX 11-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.



PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145374.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147182.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 25-AUG-1999; 99US-0150864.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0152370.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.



QY	236	ATGGAATGCAAAATTAAGATCGAATAGCGACGACGCGCAAAACCTGTGAAGACTTAATGGC	295
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	3706	DANDRAMTIVKSTIYATNHTTRDRSARCANANDSTRDVMMNTTBYYHRA-TNATATTN	3764
QY	296	TGGGTGGAATCGGCCCTAATGTCTCCAGATGACCTGGAGGAGAAGTTCACAGACTTTG	355
Db	3765	YSGRTTHNACKWSNSTTTATNTANDPTRADMAKATTTNTNSATRA-----AAGAGNMAN	3812
QY	356	GGAAATCGAGATTTTACGTTTCTCAGAGAAACGACAGACTTTCATGATTTATG	415
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	3820	MGRRAATMBSHMINDSKDCTMDMYKKDADACDMNTGTHYCHNNRBBHDHDDNDRBVB	3879
QY	416	AGATGATGATGCTTTACAGCTAAGACATGAATGAAAGCTTGGGTGTACTTT	475
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	3880	RVHDDGVMMNNHVMYDHDHDDHDDHNDHNDHCNNHBRVHDHDDVYMMNGHMHNN	3933
QY	476	TGCGTGTGATTTTCTCGGGTCACAAAGCGCCAAAAAAGAACATGGCTGGCTTACG	535
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	3940	HNHNTICDHDHDDHNTCHABDGRMYVVVVVVVVVVVVVYVCKHNDHDKHGMCHCCM	3999
QY	536	ATAACGAGATGCGCAATATGATCATTAACCGCAGTATCTCAGTATGAGACTTTA	595
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	4000	HKMDRAAANRACNNHHNADGCKHADATARA-----AHTHADRTRRAGNMER	4052
QY	596	AAGGAGATGTCACGCAAGTAAGTTCGTGGATGGGTGCGCTCTAGCTGACACAAT	655
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	4053	NNMBRSSCDMCKKMMKKVKNBDVRYTTGGBDRKNNNDYSMBSHMINDSKDQNDYS	4112
QY	656	GGAATGATGAGCAATTTCTGCACATTCGCGATGATCTTTGGTGAGATCAGAGGGTAA	715
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	4113	CTMATRCVRYSAVASDMDNDYSCTMATRCVRYSAVASDMDNYSCTMATRCVRYSAVASDMN	4172
QY	716	AAAGTTACCATCAAGGAATTTTGCACCTTGAGAGTTAGAGCGCGGAGAGAACTGCC	775
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	4173	DYSCMATRCVRYSAVASDMDNDYSCTMATRCVRYSAVASDMDNDYSCMATRCVRYSAVAS	4232
QY	776	AATGCAAGGAAGCCTACAGGGAGGCTTATTCATTAATCTAGAAATCAAAATTAATGACT	835
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	4233	DRKRSWMSKMSMCTMNMWMANSYMBAFRCRNMHHBDSDBHBSSTDWDDCSTMAT5	4292
QY	836	CAAAAGATGAGTGGCTCTGAGCAAGACGACTACTAGTTTACTTCGGTAT	887
		:::   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :	
Db	4293	SCDTWMDYSWMDYSHHSSWMDYSNDNDSCHMDYSDMCRSTWRDRSDGRTIH	4344
RESULT 9			
ID	AA58705	standard; cDNA; 256 bp.	
XX	AC	AA58705;	
XX	DT	13-FEB-2002 (first entry)	
DE	CDNA #1381	encoding portion of a human colon tumour protein.	
OS	Human;	colon tumour protein; colon cancer; gene therapy; cytostatic; ss.	
XX	Homo sapiens.		
XX	PN	WO200173027-A2.	
XX	PD	04-OCT-2001.	
XX	PF	22-MAR-2001; 2001WO-US09246.	
XX	PR	24-MAR-2000; 2000US-191597P.	
XX	PR	04-MAY-2000; 2000US-202024P.	
XX	PR	05-MAY-2000; 2000US-202189P.	
XX	PA	(CORI-) CORIXA CORP.	
XX	PI	Meagher MJ, Xu J, King GE;	
XX			

DR WPI; 2001-611627/70.

XX New colon tumour proteins and related nucleic acid, useful for  
PT treatment, prevention, diagnosis and monitoring of cancer -  
XX  
PS  
XX

Claim 4; Page 273; 299pp; English.

CC Th present invention relates to the isolation of novel cDNA sequences  
CC encoding for at least an immunogenic portion of human colon tumour  
CC proteins. The sequences of the invention are useful in pharmaceutical  
CC compositions and vaccines for the prevention and treatment of cancers  
CC such as colon cancer. They are also useful for the diagnosis and  
CC monitoring of such cancers. Antibodies to the colon tumour proteins  
CC and antigen presenting cells that express polynucleotides encoding  
CC colon tumour proteins can be used to inhibit the development of  
CC cancers. T-cells that react specifically with colon tumour proteins  
CC are useful for removing tumour cells from samples (e.g. blood) and  
CC for cancer treatment. The polynucleotide sequences are also useful in  
CC gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the  
CC invention that encode for portions of human colon tumour proteins.  
XX  
XX

Sequence 256 BP; 67 A; 46 C; 83 G; 56 T; 4 other;

	Query Match	1.6%	Score 42.8	DB 23	Length 256
	Best Local Similarity	59.7%	Pred. No. 0.011		
	Matches	71	Conservative	0	Mismatches 48
				Indels	0
				Gaps	0
Qy	151	TCTTCAGSGGTTTCCGTTTATATACACACATGTGGAGGAGACAGTCGCACCCAA	210		
Db	89	TCTCCAAAGGTTTGTTGTTTATACCTTCAACAGTCAGGAGGTGCGCAAGTGGCAAN	148		
Qy	211	GAGGCTCTTCAGAGGACCAATTGATGGAAGTCAATTAAGATCGACGACGCC	269		
Db	149	GAGGCGCATGAGACGCTGTAATTGATGGAATTAAGTGAACCTTGAGCTGGGCGCAACC	207		

Accession	Source	Organism	Gene	Protein	Function	Reference
AB199579	standard	CDNA	1595 BP			
AB199579	standard	CDNA	1595 BP			
AB199579	standard	CDNA	1595 BP			
07-MAR-2002	(first entry)					
Mouse	ischaemic condition related	cdna	sequence	SEQ ID NO:584		
Mouse	ischaemia; compressive	ischaemia; occlusive	ischaemia; vasospastic	ischaemia; ischaemic condition; ischaemic disease; ss		
Mus	musculus					
MO200188188-A2						
22-NOV-2001						
18-MAY-2001	2001WO-0P04192					
18-MAY-2000	2000JP-0145977					
(UYNF-)	UNIV NIHON SCHOOL JURIDICAL PERSON					
Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y						
WPI: 2002-034733/04						
P-PSDB: ABR57226						
Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes						
Claim 2; Page 1517-1520; 2690pp; English						



PT effects by determining the changes in gene expression in tissues or  
 PT cells exposed to the toxin and comparing these to gene expression in  
 PT unexposed tissues or cells -

PS Claim 1; Seq ID No 1567; 239pp; English.

CC The invention relates to methods for predicting toxic effects of  
 CC compounds or the progression of these toxic effects by determining the  
 CC global changes in gene expression in tissues or cells exposed to the  
 CC toxin and comparing these to gene expression in unexposed tissues or  
 CC cells. Also included are methods of predicting at least one toxic  
 CC effect of a compound or progression of a toxic effect, preferably the  
 CC hepatotoxicity of a compound, comprising detecting the level of  
 CC expression in a tissue or cell sample exposed to the compound of two or  
 CC more genes listed in the specification, where differential expression of  
 CC the genes is indicative of at least one toxic effect or progression.  
 CC The method can also be used to identify an agent which modulates the  
 CC toxic response and predict cellular pathways that a compound modulates the  
 CC support in kit form), where each of at least two probes (on a solid  
 CC specifically hybridises to a gene listed in the specification, a computer  
 CC system comprising a database containing information identifying the  
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
 CC set of genes comprising at least two genes listed in the specification,  
 CC and a user interface to view the information used to present information,  
 CC identifying the expression level in a tissue or cell of at least one gene  
 CC listed in the specification. The method is useful for elucidating global  
 CC changes in gene expression and for identifying toxicity markers in  
 CC tissues or cell exposed to a known toxin. The genes may be used as  
 CC toxicity markers in drug screening and toxicity assays. The genes and  
 CC gene expression information may be used as diagnostic markers for the  
 CC prediction or identification of the physiological state of tissue or cell  
 CC sample that has been exposed to a compound or agent. Hepatotoxicity  
 CC is characterised by centrilobular necrosis and steatosis. The present  
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
 CC which is differentially expressed in response to a hepatotoxic agent.

SX Sequence 2142 BP; 760 A; 347 C; 620 G; 415 T; 0 other;

Query Match 1.5%; Score 41.6; DB 24; Length 2142;  
 Best Local Similarity 59.28; Pred. No. 0.11;  
 Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 150 TTCTTCACGAGCTTGGCTTTATATACACATGTGGAGAGAGTCGACACCAA 209  
 Db 1830 TTCTTCACGAGCTTGGCTTTATATACACATGTGGAGAGAGTCGACACCAA 1889  
 QY 210 AGAGGCTCTTCACGAGCAATTTGAATGAGTCAATTTAGATGAGTACGACGACC 269  
 Db 1890 GGAGGCCATGAGAGATGAGAAATTTGATGAGAAACAAAGTTACTGCGGCGCAAC 1949

RESULT 13

AB199466  
 ID AB199466 standard; cDNA; 2052 BP.

AC AB199466;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:417.

KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN WO200188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

DR WPI: 2002-034733/04.

PR P-PSDB: ABB57172.

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or  
 PT by determining the expression profile of a gene group comprising these  
 PT genes -

PS Claim 2; Page 1135-1137; 2690pp; English.

CC The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (AB199202 to AB199912, encoding  
 CC the protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The  
 CC expression levels or expression profiles produced by these genes are  
 CC used as an indicator when screening for ischaemic condition improving  
 CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914  
 CC represent PCR primers for a mouse ischaemic condition related sequence,  
 CC which are used in the exemplification of the present invention.

SX Sequence 2052 BP; 589 A; 400 C; 483 G; 580 T; 0 other;

Query Match 1.5%; Score 41.4; DB 24; Length 2052;  
 Best Local Similarity 57.38; Pred. No. 0.12;  
 Matches 75; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 153 TTCACGAGCTTGGCTTTATATACACATGTGGAGAGAGTCGACACCAA 212  
 Db 616 TTCACGAGATTTGGCTTTGATATTTGAAATGTAGACATGCGAAGAGCTAAAGA 675  
 QY 213 GGCTCTTCACGAGCAATTTGAATGAGTCAATTTAGATGAGTACGACGCGC 272  
 Db 676 ACCTGCCAATGGAATGAGACCTTATGAGGCGCTGAAATTGAGTCTCTATACAAA 735  
 QY 273 AAAACCTTETA 283  
 Db 736 AAGGCCCATTA 746

RESULT 14

AACT6700  
 ID AACT6700 standard; cDNA; 11680 BP.

AC AACT6700;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2255 polynucleotide sequence SEQ ID NO:4509.

KW Human; open reading frame; ORFX; defection; cytosolic; hepatotropic;  
 KW vulnerability; antiparietal; antiparkinsonian; neurotrophic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiallergic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antihemetic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;

```

RESULF 15
ABL33124/c
ID ABL33124 standard; DNA; 6104 BP.
XX
XX
AC ABL33124;
XX
XX
DT 26-MAR-2002 (first entry)
XX
XX
Human immune system associated gene SEQ ID NO: 1097.
DE
XX
Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiatherosclerotic; antianaemic; cytosinatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflamatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A. Piepenbrock C, Berlin K.
XX
XX
WP: 2002-130909/17.
XX
XX
Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX
PS Claim 1; SEQ ID NO 1097; 32pp + Sequence listing; German.
XX
XX
The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/intestine bowel
CC diseases. The present sequence is a gene of the invention.
XX
XX
Sequence 6104 BP; 1398 A; 129 C; 1826 G; 2751 T; 0 other;
SQ
Query Match 1.5%; Score 39.6; DB 24; Length 6104;
Best Local Similarity 51.7%; Pred. No. 0.89;
Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
OY 2359 CAATACCAACAGACGCTCCATGCAACCAAACTATTGGCCAAATTCAGTTATCAG 2418
DB 3450 CATTAAAAAATAAATTTCTTACTTAACCAATCTTAACCCACGCAATTTTAAAC 3391
OY 2419 CAAGCTAATTTTCATGGCGTAACAACAATCAGGACAGAACTTAAACCTTCCCAATTT 2478
DB 3390 CACCTATATCTACATTAATAAATAAATAAATTTTAAACAAAAAATCCCTACCTAATTT 3331
OY 2479 CAACCTGCGAATGCAACCAACGACGATTAAGCAATTTAAGCCACAAACCA 2532
DB 3330 CACGACTCATTCATCTACCTCTTAACCAACCAACCTCTCAACACAAAAA 3277

```

**This Page Blank (uspto)**